

L Number	Hits	Search Text	DB	Time stamp
7	1	wo NEAR "9605306"	USPAT; US-PGPUB; EPO; JPO; DERWENT	2003/06/27 16:52
-	2	("5654155").PN.	USPAT; US-PGPUB; EPO; JPO; DERWENT	2003/06/23 13:00
-	1	("20020183268").PN.	USPAT; US-PGPUB; EPO; JPO; DERWENT	2003/01/14 17:03
-	9	(Murphy NEAR Patricia) and BRCA1	USPAT; US-PGPUB; EPO; JPO; DERWENT	2003/01/14 17:07
-	170	BRCA1 WITH sequence	USPAT; US-PGPUB; EPO; JPO; DERWENT	2003/01/14 17:09
-	164	(BRCA1 WITH sequence) and gene	USPAT; US-PGPUB; EPO; JPO; DERWENT	2003/01/14 17:08
-	12	((BRCA1 WITH sequence) and gene) and omi\$1	USPAT; US-PGPUB; EPO; JPO; DERWENT	2003/01/14 17:08
-	8	BRCA1 SAME sequence SAME gene SAME omi\$1	USPAT; US-PGPUB; EPO; JPO; DERWENT	2003/01/14 17:09
-	116	Stommel	USPAT; US-PGPUB; EPO; JPO; DERWENT	2003/06/02 14:14
-	1	Stommel and brca1	USPAT; US-PGPUB; EPO; JPO; DERWENT	2003/06/02 14:15
-	700	Durocher	USPAT; US-PGPUB; EPO; JPO; DERWENT	2003/06/02 14:15
-	10	Durocher and brca1	USPAT; US-PGPUB; EPO; JPO; DERWENT	2003/06/02 14:15
-	20	holt NEAR jeffrey	USPAT; US-PGPUB; EPO; JPO; DERWENT	2003/06/18 12:41
-	11	(holt NEAR jeffrey) and brca1	USPAT; US-PGPUB; EPO; JPO; DERWENT	2003/06/18 12:41
-	962	BRCA1	USPAT; US-PGPUB; EPO; JPO; DERWENT	2003/06/18 12:51
-	13	BRCA1 and (thymidine WITH cytidine)	USPAT; US-PGPUB; EPO; JPO; DERWENT	2003/06/18 12:52
-	28	Murphy NEAR Patricia	USPAT; US-PGPUB; EPO; JPO; DERWENT	2003/06/19 10:46



1  
2  
3  
4  
5



-	6	(Murphy NEAR Patricia) and SEQ.clm.	USPAT; US-PGPUB; EPO; JPO; DERWENT	2003/06/19 10:46
-	49	((("5547839") or ("5510270") or ("6045997") or ("5561058") or ("5455934") or ("5891857") or ("6051379") or ("5858669") or ("4683202") or ("6130322") or ("5750400") or ("5911227") or ("5624803") or ("6083698") or ("5545531") or ("5948643") or ("5693473") or ("5589330") or ("5633134") or ("5726019") or ("5710001") or ("5753441") or ("5747282") or ("6033857") or ("6124104") or ("5756294"))).PN.	USPAT; US-PGPUB; EPO; JPO; DERWENT	2003/06/23 13:15
-	2	("5912127").PN.	USPAT; US-PGPUB; EPO; JPO; DERWENT	2003/06/23 13:16



Db 301 GTCCCTTATGTAAGATGATTAATACCAAAAGAGCCTACAAGAAATGACAGATTTCAGTC 360  
OY 361 AACTGTGTAAGACCTAATGAAATCATTTTGTGCTTTTCAGCTTGAACAGGTTTGAGT 420  
Db 361 AACTGTGTAAGACCTAATGAAATCATTTTGTGCTTTTCAGCTTGAACAGGTTTGAGT 420  
OY 421 ATGCAACAGCTAATATTTTGCAGAAAAGAAAATAACTCTCCGGAACATCTAAAGATG 480  
Db 421 ATGCAACAGCTAATATTTTGCAGAAAAGAAAATAACTCTCCGGAACATCTAAAGATG 480  
OY 481 AAGTTCTATCATCAAGATATGAGGCTACAGAAACGCTGCCAAAAGACTTCTACAGAGT 540  
Db 481 AAGTTCTATCATCAAGATATGAGGCTACAGAAACGCTGCCAAAAGACTTCTACAGAGT 540  
OY 541 AACCAGAAATCCTCTCTGAGGAGAACCACTCATGTCCTCAACCTCTACCTTGGA 600  
Db 541 AACCAGAAATCCTCTCTGAGGAGAACCACTCATGTCCTCAACCTCTACCTTGGA 600  
OY 601 CTGTGAGAACTCTGAGAGCAAGAGCGAGTACAACTCAAAAGAGCTGTCTACATG 660  
Db 601 CTGTGAGAACTCTGAGAGCAAGAGCGAGTACAACTCAAAAGAGCTGTCTACATG 660  
OY 661 AATTGGATCTGATCTTCTGAGAGTACCGTTAATAAGCACTTATTCAGTGTGGAG 720  
Db 661 AATTGGATCTGATCTTCTGAGAGTACCGTTAATAAGCACTTATTCAGTGTGGAG 720  
OY 721 ATCAAGAAATGTTACAAATACACCCCTCAGAGAACCGAGTGAATAGTTGGATTCTG 780  
Db 721 ATCAAGAAATGTTACAAATACACCCCTCAGAGAACCGAGTGAATAGTTGGATTCTG 780  
OY 781 CAAAAAGGCTGCTTGTGAATTTTCTGAGAGCGAGTGTACAAATACTGAACATCATAC 840  
Db 781 CAAAAAGGCTGCTTGTGAATTTTCTGAGAGCGAGTGTACAAATACTGAACATCATAC 840  
OY 841 CCACTAATTAATGATTTGACACACACAGAGCGTGAAGGCTCAGAGGCTCAGAAAAGT 900  
Db 841 CCACTAATTAATGATTTGACACACACAGAGCGTGAAGGCTCAGAGGCTCAGAAAAGT 900  
OY 901 ATCAGGATGATCTGTTCAAACTGATGATGAGGCGATGAGCAAAATACATGAGCA 960  
Db 901 ATCAGGATGATCTGTTCAAACTGATGATGAGGCGATGAGCAAAATACATGAGCA 960  
OY 961 GCTCATTACAGCATGAGAACAGCAGTTTATTACTCATAAAGACAGAAATGATGAAA 1020  
Db 961 GCTCATTACAGCATGAGAACAGCAGTTTATTACTCATAAAGACAGAAATGATGAAA 1020  
OY 1021 AGGCTGAATCTGTAATAAAGCAAAAGCGCTGCTTACAGAGAGCAACATTAACAGAT 1080  
Db 1021 AGGCTGAATCTGTAATAAAGCAAAAGCGCTGCTTACAGAGAGCAACATTAACAGAT 1080  
OY 1081 GGGCTGGAATGAAGAAACATGTAATGATAGGCGGAGCTCCAGACAGAAAAGGTTAG 1140  
Db 1081 GGGCTGGAATGAAGAAACATGTAATGATAGGCGGAGCTCCAGACAGAAAAGGTTAG 1140  
OY 1141 ATCTGAATGCTGATCCCTGTGTGAGAGAAAAGAAATGGAATAGCAGAAATCCATGCT 1200  
Db 1141 ATCTGAATGCTGATCCCTGTGTGAGAGAAAAGAAATGGAATAGCAGAAATCCATGCT 1200  
OY 1201 CAGAGAACTCTAGAGATGCTGAGAGATGCTCTGATACACTTAATAGCAGATTTCAGA 1260  
Db 1201 CAGAGAACTCTAGAGATGCTGAGAGATGCTCTGATACACTTAATAGCAGATTTCAGA 1260  
OY 1261 AAGTTAATGAGTGTTCAGAGAGTGAATGAGTGTAGTGTCTGATGACTCAGATGAG 1320  
Db 1261 AAGTTAATGAGTGTTCAGAGAGTGAATGAGTGTAGTGTCTGATGACTCAGATGAG 1320  
OY 1321 GGGAGTCTGAATCAAAATGCAAAAGTACGATGATTTGACAGCTTCAAAATGAGTAGAG 1380  
Db 1321 GGGAGTCTGAATCAAAATGCAAAAGTACGATGATTTGACAGCTTCAAAATGAGTAGAG 1380  
OY 1381 AATATTCTGCTCTTCTGAGAGAAATAGACTTACTGGCCAGTGATCTCTCATGAGGCTTAA 1440  
Db 1381 AATATTCTGCTCTTCTGAGAGAAATAGACTTACTGGCCAGTGATCTCTCATGAGGCTTAA 1440

OY 1441 TATGTAAAGTGAAGAGTTCACCTCAATCAGTAGAGAGTAATATTGAGACAAATAT 1500  
Db 1441 TATGTAAAGTGAAGAGTTCACCTCAATCAGTAGAGAGTAATATTGAGACAAATAT 1500  
OY 1501 TTGGGAAACCTATGCGAAGAGGCAAGCCTCCCAACTTAAGCAGTATTAAGTGAATC 1560  
Db 1501 TTGGGAAACCTATGCGAAGAGGCAAGCCTCCCAACTTAAGCAGTATTAAGTGAATC 1560  
OY 1561 TAAATATGAGAGATTTGTTACTGAGCCACAGATTAATACAGAGCGTCCCTCAGAAAT 1620  
Db 1561 TAAATATGAGAGATTTGTTACTGAGCCACAGATTAATACAGAGCGTCCCTCAGAAAT 1620  
OY 1621 AATTAAGGCTAAAGAGAGCTACATCAGAGCCTTCATCCTGAGATTTATATCAAGAG 1680  
Db 1621 AATTAAGGCTAAAGAGAGCTACATCAGAGCCTTCATCCTGAGATTTATATCAAGAG 1680  
OY 1681 CAGATTTGAGAGTTCAGAAACACCTCGAATGATTAATACAGGAGACTTAACCAAGAGC 1740  
Db 1681 CAGATTTGAGAGTTCAGAAACACCTCGAATGATTAATACAGGAGACTTAACCAAGAGC 1740  
OY 1741 AGAATGCTCAAGTATGATTAATTAATTAATAGTGTCTAGAGATTAACCAAGAGTAT 1800  
Db 1741 AGAATGCTCAAGTATGATTAATTAATTAATAGTGTCTAGAGATTAACCAAGAGTAT 1800  
OY 1801 CFATTCAGATGAGAAATTCCTAACCCATAGATACCTCGAAGAAAGATCTGCTTCA 1860  
Db 1801 CFATTCAGATGAGAAATTCCTAACCCATAGATACCTCGAAGAAAGATCTGCTTCA 1860  
OY 1861 AAGCAAGAGCTGAACCTATTAAGCAGAGATTAAGCAATATGAGAACTGAAATTAATATC 1920  
Db 1861 AAGCAAGAGCTGAACCTATTAAGCAGAGATTAAGCAATATGAGAACTGAAATTAATATC 1920  
OY 1921 ACAATTCAGAAAGCAGCTTAAGAAATAGGCTGAGAGAGAGTCTTACAGGCAATATTC 1980  
Db 1921 ACAATTCAGAAAGCAGCTTAAGAAATAGGCTGAGAGAGAGTCTTACAGGCAATATTC 1980  
OY 1981 ATGCGCTGAGCAATGATAGTACAGTAAGTAATTAAGCCACCTTAATGATGATTCGAA 2040  
Db 1981 ATGCGCTGAGCAATGATAGTACAGTAAGTAATTAAGCCACCTTAATGATGATTCGAA 2040  
OY 2041 TTGATAGTGTCTTACAGAGTGAAGAGATTAAGAAAAAGTACCAACCAATCCAGTCA 2100  
Db 2041 TTGATAGTGTCTTACAGAGTGAAGAGATTAAGAAAAAGTACCAACCAATCCAGTCA 2100  
OY 2101 GGCACAGCAGAAACCTACCAATCTATGAAAGGTAAAGAACCTGCAACTGAGGCCAAGAA 2160  
Db 2101 GGCACAGCAGAAACCTACCAATCTATGAAAGGTAAAGAACCTGCAACTGAGGCCAAGAA 2160  
OY 2161 GTACCAAGCCCAATGAAAGAGACAGTAAGAAAGACATGACAGTGTATCTTCCAGAGCTGA 2220  
Db 2161 GTACCAAGCCCAATGAAAGAGACAGTAAGAAAGACATGACAGTGTATCTTCCAGAGCTGA 2220  
OY 2221 AGTTAACAATGACAGCTGCTTCTTACTAAGTGTTCACCAATACAGTGAATTAAGAAAT 2280  
Db 2221 AGTTAACAATGACAGCTGCTTCTTACTAAGTGTTCACCAATACAGTGAATTAAGAAAT 2280  
OY 2281 TTGTCAATCTGACCTTCCAGAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGT 2340  
Db 2281 TTGTCAATCTGACCTTCCAGAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGT 2340  
OY 2341 CTATATATGCTGAGAGAGCCCAAGAGTCTATGTTAAGTGAAGAGGTTTGGCAAGC 2400  
Db 2341 CTATATATGCTGAGAGAGCCCAAGAGTCTATGTTAAGTGAAGAGGTTTGGCAAGC 2400  
OY 2401 AAGAGATCTGAGAGAGTACAGTATTTCAATGCTGACTGCTACTGATTAATGACACTCAG 2460  
Db 2401 AAGAGATCTGAGAGAGTACAGTATTTCAATGCTGACTGCTACTGATTAATGACACTCAG 2460  
OY 2461 AAGATATCTGCTTACTGAGAGTATGAGACTCTAGGAGAGCAAAAGCAGAAACCAATTAAT 2520  
Db 2461 AAGATATCTGCTTACTGAGAGTATGAGACTCTAGGAGAGCAAAAGCAGAAACCAATTAAT 2520

Db	4921	AAGTTCCCAATTTGAAAGTTCAGAAATCTGGCCAGAGTCCAGCTGCTCATCTACTG	4980
QY	4981	ATACTGCTGGGTATATATGCAATTCGAAAGAGTGTGACGAGGAGAGACCAATTTGACG	5040
Db	4981	ATACTGCTGGGTATATATGCAATTCGAAAGAGTGTGACGAGGAGAGACCAATTTGACG	5040
QY	5041	CTTCAACGAAAGGGTCAACAAAGAAATGTCATGCTGCTGCTGCGCTGACCCAGAG	5100
Db	5041	CTTCAACGAAAGGGTCAACAAAGAAATGTCATGCTGCTGCTGCGCTGACCCAGAG	5100
QY	5101	AATTTATCTGCTGTACAAAGTTTGGCCAGAAACACACATCATCTTAACTATCTAATTA	5160
Db	5101	AATTTATCTGCTGTACAAAGTTTGGCCAGAAACACACATCATCTTAACTATCTAATTA	5160
QY	5161	CTGAAGACACTACTCATGTTGTTATGAAAGACAGATGCTGAATTTGTGTGTAAGGACAC	5220
Db	5161	CTGAAGACACTACTCATGTTGTTATGAAAGACAGATGCTGAATTTGTGTGTAAGGACAC	5220
QY	5221	TGAAATATTTTCTAGAAATTCGGGAGAGAAATGGGTAGTTAGCTATTTCTGGGTGACC	5280
Db	5221	TGAAATATTTTCTAGAAATTCGGGAGAGAAATGGGTAGTTAGCTATTTCTGGGTGACC	5280
QY	5281	AGTCATTTAAAGAAAGAAATTCGTAATGATGATGATTTGAAGTCAGAGACAGATGGG	5340
Db	5281	AGTCATTTAAAGAAAGAAATTCGTAATGATGATGATTTGAAGTCAGAGACAGATGGG	5340
QY	5341	TCAATGGAAGAAACCCACCAGGCTCCAAAGCAGAGAGAAATTCACAGACAGAAAGATCT	5400
Db	5341	TCAATGGAAGAAACCCACCAGGCTCCAAAGCAGAGAGAAATTCACAGACAGAAAGATCT	5400
QY	5401	TCAGGGGGCTTGAAATCTGTTGCTATGAGGCCCTTCCACCACATGCCACAGATCAACTGG	5460
Db	5401	TCAGGGGGCTTGAAATCTGTTGCTATGAGGCCCTTCCACCACATGCCACAGATCAACTGG	5460
QY	5461	AATGAGTGTACAGCTGTGTGCTGCTCTCTGTGTGTAAGAGACTTTCATCATTCACCTTG	5520
Db	5461	AATGAGTGTACAGCTGTGTGCTGCTCTCTGTGTGTAAGAGACTTTCATCATTCACCTTG	5520
QY	5521	GCACAGGTGTCCACCCCAATTTGGTGTGCGACAGAGATGCTTGACAGAGACAAATGGCT	5580
Db	5521	GCACAGGTGTGTCCACCCCAATTTGGTGTGCGACAGAGATGCTTGACAGAGACAAATGGCT	5580
QY	5581	TCCATGCAATTTGGGCAGATGTGTAGAGCACCTGTGTGACCCGAGAGTGGGTGTTGGACA	5640
Db	5581	TCCATGCAATTTGGGCAGATGTGTAGAGCACCTGTGTGACCCGAGAGTGGGTGTTGGACA	5640
QY	5641	GTGTAGCACTTACCAGTGGCCAGAGAGCTGGACACCTTACCTGATATACCCAGATCCCCACA	5700
Db	5641	GTGTAGCACTTACCAGTGGCCAGAGAGCTGGACACCTTACCTGATATACCCAGATCCCCACA	5700
QY	5701	GCCACTACTGA 5711	
Db	5701	GCCACTACTGA 5711	

XX	RESULT 6
XX	AAV46464
ID	AAV46464 standard; cDNA: 5711 BP.
XX	AAV46464;
AC	
XX	
DT	18-NOV-1998 (first entry)
XX	
DE	Human BRCA1 om12 polymorphism #7 cDNA.
XX	
XX	BRCA1, om1; human; breast and ovarian cancer predisposing gene;
KW	polymorphism; susceptibility; anti-oncogene; tumour suppressor;
KW	chromosome 17q; ss.
XX	
OS	Homo sapiens.
XX	
XX	
Key	Location/Qualifiers
PH	

FT	CDS	120..5711
FT		/tag= a
FT		/product= "BRCA1 omi2 protein"
FT		4956
FT	variation	/tag= b
FT		/note= "This polymorphic variation can be an A or G nucleotide"
XX		
XX		
PN	US5750400-A.	
PD	12-MAY-1998.	
XX		
XX	12-FEB-1997;	97US-0798691.
PF	12-FEB-1996;	96US-0598591.
PR	12-FEB-1997;	97US-0798691.
XX		
PA	(ONCO-) ONCORMED INC.	
PI	Allen AC, Alvares CP, Critz BS, Murphy PD, Olson SJ;	
PI	Scheller DB, Zeng B;	
DR	WPI; 1998-296774/26.	
XX		
PT	BRCA1 omi gene coding sequences - useful for distinguishing between	
PT	polymorphisms and mutation(s) in the screening for disposition to	
PT	breast or ovarian cancer	
PS	Claim 2e; Page -: 54pp; English.	
CC	This sequence encodes a human BRCA1 (breast and ovarian cancer	
CC	predisposing gene) omi2 gene in which a polymorphic variation occurs at	
CC	nucleotide 4956. This sequence and other polymorphic variations of this	
CC	sequence are useful for the identification of an individual who may or	
CC	may not have an increased susceptibility to breast or ovarian cancer.	
CC	The sequences used identify gene changes which are due to polymorphisms,	
CC	rather than cancer causing mutations. BRCA1 is an anti-oncogene (tumour	
CC	suppressor) which is involved in genetic inheritance of cancers,	
CC	especially breast and ovarian cancer. It is found at human chromosome	
CC	17q which is known to be linked to cancer susceptibility, especially	
CC	breast cancer. Cells containing a mutation in this gene lose the	
CC	wild-type function of BRCA1 and are more susceptible to cancers.	
CC	NOTE: This sequence does not appear in the specification but has been	
CC	created from the wild type BRCA1 omi2 gene represented in AAU46449.	
XX		
SO	Sequence 5711 BP; 1955 A; 1098 C; 1274 G; 1383 T; 1 other;	
Query Match	100.0%; Score 5709; DB 19; Length 5711;	
Best Local Similarity	100.0%; Pred. No. 0;	
Matches 5709; Conservative	1; Mismatches	1; Indels 0; Gaps 0;
OY	1 AGCTCGCTGAGACTTTCCTGGACCCCGCACAGGCTGTGGGTTTCTCAGATACTGGGCC	60
Db		
Db	1 AGCTCGCTGAGACTTTCCTGGACCCCGCACAGGCTGTGGGTTTCTCAGATACTGGGCC	60
OY	61 CCTGCGCTCAGAGGCTTACCCTCTGCTGTGGGTAAAGTTCAATTGGACAGAAAAGA	120
Db	61 CCTGCGCTCAGAGGCTTACCCTCTGCTGTGGGTAAAGTTCAATTGGACAGAAAAGA	120
OY	121 TGGATTTATCTGCCTCTTGCGCTTGAAGAAGTACAAAGTCAATTATGCTTGCAGAAAA	180
Db	121 TGGATTTATCTGCCTCTTGCGCTTGAAGAAGTACAAAGTCAATTATGCTTGCAGAAAA	180
OY	181 TCCTTAGAGTGCCTCATCTGCTGGAGTTGATCAAGAACCTGTCCCAACAAGTGTGACC	240
Db	181 TCCTTAGAGTGCCTCATCTGCTGGAGTTGATCAAGAACCTGTCCCAACAAGTGTGACC	240
OY	241 ACAATATTTGGCAAATTTTGGATGCTGGAACCTTCCAACCAAGAAAGGCGCTTACAGT	300
Db	241 ACAATATTTGGCAAATTTTGGATGCTGGAACCTTCCAACCAAGAAAGGCGCTTACAGT	300
OY	301 GTCTTTATGTAAGATGATTAACCAAAAGAGCCTPACAGAAAGTACGAGATTTAGTC	360
Db		

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 26, 2003, 12:34:02 ; Search time 14215.3 Seconds

(without alignments)  
11692.029 Million cell updates/sec

Title: US-09-734-672-3

Perfect score: 5711

Sequence: 1 AGCTCGCTGACACTTCTGCG.....TCCCCCAGCGCACTACTGA 5711

Scoring table: IDENTITY\_NUC

Gapop 10.0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: \*  
1: gb\_ba: \*  
2: gb\_hlg: \*  
3: gb\_in: \*  
4: gb\_cm: \*  
5: gb\_ov: \*  
6: gb\_pat: \*  
7: gb\_ph: \*  
8: gb\_pl: \*  
9: gb\_pr: \*  
10: gb\_ro: \*  
11: gb\_sts: \*  
12: gb\_sy: \*  
13: gb\_un: \*  
14: gb\_vl: \*  
15: em\_ba: \*  
16: em\_fun: \*  
17: em\_hum: \*  
18: em\_in: \*  
19: em\_mu: \*  
20: em\_om: \*  
21: em\_or: \*  
22: em\_ov: \*  
23: em\_pat: \*  
24: em\_ph: \*  
25: em\_pl: \*  
26: em\_ro: \*  
27: em\_sts: \*  
28: em\_un: \*  
29: em\_vl: \*  
30: em\_hlg\_hum: \*  
31: em\_hlg\_inv: \*  
32: em\_hlg\_other: \*  
33: em\_hlg\_mus: \*  
34: em\_hlg\_pin: \*  
35: em\_hlg\_rod: \*  
36: em\_hlg\_mam: \*  
37: em\_hlg\_vtl: \*  
38: em\_sy: \*  
39: em\_hlgo\_hum: \*  
40: em\_hlgo\_mus: \*  
41: em\_hlgo\_other: \*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5709.4	100.0	5711	AR007334	AR007334 Sequence
2	5709.4	100.0	5711	AR112810	AR112810 Sequence
3	5707.8	99.9	5711	AR033056	AR033056 Sequence
4	5707.8	99.9	5711	HSU14680	U14680 Homo sapien
5	5707.8	99.9	5712	AR070223	AR070223 Sequence
6	5707.8	99.9	5712	AR118507	AR118507 Sequence
7	5707.8	99.9	5712	AR125601	AR125601 Sequence
8	5707.8	99.9	5712	AR184044	AR184044 Sequence
9	5707.8	99.9	5914	AR004673	AR004673 Sequence
10	5707.8	99.9	5914	AR008159	AR008159 Sequence
11	5707.8	99.9	5914	AR136942	AR136942 Sequence
12	5707.8	99.9	5914	I76943	I76943 Sequence 1
13	5707.8	99.9	5914	I80938	I80938 Sequence 1
14	5707.8	99.9	5914	I81034	I81034 Sequence 1
15	5706.2	99.9	5711	AR048660	AR048660 Sequence
16	5706.2	99.9	5711	AR048666	AR048666 Sequence
17	5706.2	99.9	5711	I40795	I40795 Sequence 4
18	5706.2	99.9	5711	I40801	I40801 Sequence 10
19	5704.6	99.9	5711	AR007335	AR007335 Sequence
20	5704.6	99.9	5711	AR112809	AR112809 Sequence
21	5703	99.9	5711	AR007333	AR007333 Sequence
22	5703	99.9	5711	AR112808	AR112808 Sequence
23	5703	99.9	5711	I59546	I59546 Sequence 1
24	5696.8	99.8	5712	AR048668	AR048668 Sequence 12
25	5695.8	99.7	5710	AR048662	AR048662 Sequence 6
26	5695.8	99.7	5710	I40797	I40797 Sequence 6
27	5693.8	99.7	5709	AR048658	AR048658 Sequence
28	5693.8	99.7	5709	AR048663	AR048663 Sequence
29	5693.8	99.7	5709	AR048664	AR048664 Sequence
30	5693.8	99.7	5709	AR048665	AR048665 Sequence
31	5693.8	99.7	5709	I40793	I40793 Sequence 2
32	5693.8	99.7	5709	I40793	I40793 Sequence 7
33	5693.8	99.7	5709	I40798	I40798 Sequence 8
34	5693.8	99.7	5709	I40799	I40799 Sequence 8
35	5693.8	99.7	5709	I40800	I40800 Sequence 9
36	5689.8	99.6	5707	AR048667	AR048667 Sequence
37	5689.8	99.6	5707	I40802	I40802 Sequence 11
38	5657.8	99.1	5709	AX281859	AX281859 Sequence
39	5653.8	99.0	5689	AR048659	AR048659 Sequence
40	5653.8	99.0	5689	I40794	I40794 Sequence 3
41	5627.6	98.5	5770	AR048661	AR048661 Sequence
42	5627.6	98.5	5770	I40796	I40796 Sequence 5
43	5588.8	97.9	5656	AR048657	AR048657 Sequence
44	5588.8	97.9	5656	I40792	I40792 Sequence 1
45	5530.8	96.8	5693	AF005068	AF005068 Homo sapi

#### ALIGNMENTS

RESULT 1  
LOCUS AR007334 5711 bp DNA linear PAT 04-DEC-1998  
DEFINITION Sequence 3 from patent US 5750400.  
ACCESSION AR007334  
VERSION AR007334.1 GI:3966818  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 5711)  
AUTHORS Murphy,P.D., Allen,A.C., Alvares,C.P., Critz,B.S., Olson,S.J.,  
Schelter,D.B. and Zeng,B.  
TITLE Coding sequences of the human BRCA1 gene  
JOURNAL Patent: US 5750400-A 3 12-MAY-1998;

FEATURES Location/Qualifiers  
source 1. 5711 /organism="unknown"  
BASE COUNT 1956 a 1098 c 1274 g 1383 t  
ORIGIN

Query Match 100.0%; Score 5709.4; DB 6; Length 5711;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 5710; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGCTCGCTGAGACTTCTCTGACCCCGACACGAGGCTGTGGGTTTCTCAGATACTGGGCC 60  
DB 1 AGCTCGCTGAGACTTCTCTGACCCCGACACGAGGCTGTGGGTTTCTCAGATACTGGGCC 60  
QY 61 CCTGGCTCAGAGAGCCCTTCAACCCTGTGCTGGGTAAGTTCAATTGGAACAGAAAGAA 120  
DB 61 CCTGGCTCAGAGAGCCCTTCAACCCTGTGCTGGGTAAGTTCAATTGGAACAGAAAGAA 120  
QY 121 TGGATTATCTGCTCTCGCGTTGAGAGATGACAAATGTCTTAATGCTATGAGAGAAA 180  
DB 121 TGGATTATCTGCTCTCGCGTTGAGAGATGACAAATGTCTTAATGCTATGAGAGAAA 180  
QY 181 TCTTAGAGTGTCCCATCTGTCTGAGATTGATCAAGAAACCTGTCTCCACAAAGTGTACC 240  
DB 181 TCTTAGAGTGTCCCATCTGTCTGAGATTGATCAAGAAACCTGTCTCCACAAAGTGTACC 240  
QY 241 ACATATTTTGCAAATTTTGCATGCTGAAACTTCTCAACACGAAAGAGGCGCTTCACAGT 300  
DB 241 ACATATTTTGCAAATTTTGCATGCTGAAACTTCTCAACACGAAAGAGGCGCTTCACAGT 300  
QY 301 GTCTTTATGTAGAATGATATACCAAAAGAGCCCTACAAAGATGACGAGATTGATGTC 360  
DB 301 GTCTTTATGTAGAATGATATACCAAAAGAGCCCTACAAAGATGACGAGATTGATGTC 360  
QY 361 AACTGTGTAAGAGCTATTTGAAAAATCATTTGTGCTTTTCAGCTTGACACAGGTTTGAGT 420  
DB 361 AACTGTGTAAGAGCTATTTGAAAAATCATTTGTGCTTTTCAGCTTGACACAGGTTTGAGT 420  
QY 421 ATGCAAAACGCTTAATTTTGCAGAAAAAGAAAAATTAACCTCTCGCAACATCTAAAGATG 480  
DB 421 ATGCAAAACGCTTAATTTTGCAGAAAAAGAAAAATTAACCTCTCGCAACATCTAAAGATG 480  
QY 481 AAGTTTCTATCATCAAAAGTATGGGTACAGAAAACCTGCCAAAGACTTCTACAGAGTG 540  
DB 481 AAGTTTCTATCATCAAAAGTATGGGTACAGAAAACCTGCCAAAGACTTCTACAGAGTG 540  
QY 541 AACCCGAAAAATCTCTTCTGACAGAAACAGTCTCACTGTCTCAACTCTTAACCTTGAA 600  
DB 541 AACCCGAAAAATCTCTTCTGACAGAAACAGTCTCACTGTCTCAACTCTTAACCTTGAA 600  
QY 601 CTGTGAACTCTGAGAGCAAAAGCAGGAGTACAACTCAAAAGACCTGTCTACACTTG 660  
DB 601 CTGTGAACTCTGAGAGCAAAAGCAGGAGTACAACTCAAAAGACCTGTCTACACTTG 660  
QY 661 AATTTGGATCTGATTTCTTCTGAGATATACCTTAATAGGCAACTTATGAGTGGAG 720  
DB 661 AATTTGGATCTGATTTCTTCTGAGATATACCTTAATAGGCAACTTATGAGTGGAG 720  
QY 721 ATCAAGAAATTTTACAATACACCCTCAAGGAACAGGAGTGAATCAGTTTGGATCTG 780  
DB 721 ATCAAGAAATTTTACAATACACCCTCAAGGAACAGGAGTGAATCAGTTTGGATCTG 780  
QY 781 CAAAAAAGGCTGCTTGAATTTTCTGAGACGAGATGTACAAATTAAGTGAATCTATC 840  
DB 781 CAAAAAAGGCTGCTTGAATTTTCTGAGACGAGATGTACAAATTAAGTGAATCTATC 840  
QY 841 CGAGTAATATGATTTGAGACACACTGAGAGAGCTGAGAGGCAATCCAGAAAGT 900  
DB 841 CGAGTAATATGATTTGAGACACACTGAGAGAGCTGAGAGGCAATCCAGAAAGT 900  
QY 901 ATCAGAGTATGTTCTGTTTCAAACTTGATGAGGCAATGACACAAATTAATCTATG 960  
DB 901 ATCAGAGTATGTTCTGTTTCAAACTTGATGAGGCAATGACACAAATTAATCTATG 960

QY 961 GCTATTAACAGATGAGACACGACTTTATTAATCTACTATAAGACAGATGATGAGAA 1020  
DB 961 GCTATTAACAGATGAGACACGACTTTATTAATCTACTATAAGACAGATGATGAGAA 1020  
QY 1021 AGGCTGAATCTGTAATTAAGCAACAGGCTGGTTTACGAGAGCCCAATTAACAGAT 1080  
DB 1021 AGGCTGAATCTGTAATTAAGCAACAGGCTGGTTTACGAGAGCCCAATTAACAGAT 1080  
QY 1081 GGGCTGGAAGTAAAGAAACATGTAATGATGGGGAGCTCCACACAGAAAAAGGTTAG 1140  
DB 1081 GGGCTGGAAGTAAAGAAACATGTAATGATGGGGAGCTCCACACAGAAAAAGGTTAG 1140  
QY 1141 ATCTGAATGCTGATCCCTCTGTGAGAGAAAGAAATGAAATGACAGAAACCTGCT 1200  
DB 1141 ATCTGAATGCTGATCCCTCTGTGAGAGAAAGAAATGAAATGACAGAAACCTGCT 1200  
QY 1201 CAGAGAACTCAGAGATGATGAGATGCTTGGATTAACACTAAATTAACAGCATTCAGA 1260  
DB 1201 CAGAGAACTCAGAGATGATGAGATGCTTGGATTAACACTAAATTAACAGCATTCAGA 1260  
QY 1261 AAGTTAATGAGTGTGTTTCCAGAAAGTATGATGATGATGATGATGATGATGATGATG 1320  
DB 1261 AAGTTAATGAGTGTGTTTCCAGAAAGTATGATGATGATGATGATGATGATGATGATG 1320  
QY 1321 GGGAGTGTGATCAATCAATGCCAAGTATGATGATGATGATGATGATGATGATGATG 1380  
DB 1321 GGGAGTGTGATCAATCAATGCCAAGTATGATGATGATGATGATGATGATGATGATG 1380  
QY 1381 AATATCTGCTGCTTCCAGAAATTAAGTATGATGATGATGATGATGATGATGATGATG 1440  
DB 1381 AATATCTGCTGCTTCCAGAAATTAAGTATGATGATGATGATGATGATGATGATGATG 1440  
QY 1441 TATGTAAAGTGAAGAGTTTCACTCCAAATCACTAAGAGATATATGTAAGACAAATAT 1500  
DB 1441 TATGTAAAGTGAAGAGTTTCACTCCAAATCACTAAGAGATATATGTAAGACAAATAT 1500  
QY 1501 TTGGGAAAACCTATCGGAAGAGCAAGCCTCCCACTTAAGCCATGTAAGTGAATATC 1560  
DB 1501 TTGGGAAAACCTATCGGAAGAGCAAGCCTCCCACTTAAGCCATGTAAGTGAATATC 1560  
QY 1561 TAATTAATGAGAGATTTGTTACTGAGCAGATATACAAAGGCTGCCCTCACAAATA 1620  
DB 1561 TAATTAATGAGAGATTTGTTACTGAGCAGATATACAAAGGCTGCCCTCACAAATA 1620  
QY 1621 AATTAAGCCTAAAGAGAGACCTACATCAGGCTTCTATCTGAGATTTTATCAAGAAAG 1680  
DB 1621 AATTAAGCCTAAAGAGAGACCTACATCAGGCTTCTATCTGAGATTTTATCAAGAAAG 1680  
QY 1681 CAGATTTGGCAGTTCAAAAGACCTCTGAAATGATTAATCAGGGAACTAAACAGGAGC 1740  
DB 1681 CAGATTTGGCAGTTCAAAAGACCTCTGAAATGATTAATCAGGGAACTAAACAGGAGC 1740  
QY 1741 AGAATGCTCAAGTGAATGATTAATGATGATGATGATGATGATGATGATGATGATG 1800  
DB 1741 AGAATGCTCAAGTGAATGATTAATGATGATGATGATGATGATGATGATGATGATG 1800  
QY 1801 CTATTCAGATGAGAAAAATCTTAACCAATTAATTAATTAATTAATTAATTAATTAAT 1860  
DB 1801 CTATTCAGATGAGAAAAATCTTAACCAATTAATTAATTAATTAATTAATTAATTAAT 1860  
QY 1861 AAAGGAAAGCTGAACTATTAAGCAGAGATTAAGCAATATGGAATCGAATTAATATCC 1920  
DB 1861 AAAGGAAAGCTGAACTATTAAGCAGAGATTAAGCAATATGGAATCGAATTAATATCC 1920  
QY 1921 ACAATTCAAAGACCTTAATAAAGAAATGAGGTGAGGAGAAATCTTCTACAGGCAATATC 1980  
DB 1921 ACAATTCAAAGACCTTAATAAAGAAATGAGGTGAGGAGAAATCTTCTACAGGCAATATC 1980  
QY 1981 ATGCGCTTGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2040  
DB 1981 ATGCGCTTGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2040

QY	2041	TTGATAGTGGTTCTAGCAGTCGACAGATAAAGAAAAAAACTACACCAATGGCAGTC	2100
Db	2041	TTGTATAGTTGTTCTAGCAGTGAAGATAAAGAAAAAAACTACACCAATGGCAGTC	2100
QY	2101	GGCAGCAGAGAAACCTACACTCATGTGGAAGGTAAAGAACTCGCAACTGGAGCCAAAGA	2160
Db	2101	GGCAGCAGAGAAACCTACACTCATGTGGAAGGTAAAGAACTCGCAACTGGAGCCAAAGA	2160
QY	2161	GTAACAGCCAAATGAACAGACAGTAAAGACATGACAGTAATCTTTCCAGAGCTGA	2220
Db	2161	GTAACAGCCAAATGAACAGACAGTAAAGACATGACAGTAATCTTTCCAGAGCTGA	2220
QY	2221	AGTTAAACAAATGGCACCCTGGTCTTTTACTTAAGTGTCAATATCCAGTAGAATTAAAGAT	2280
Db	2221	AGTTAAACAAATGGCACCCTGGTCTTTTACTTAAGTGTCAATATCCAGTAGAATTAAAGAT	2280
QY	2281	TTGTCAATCCAGCCTTCCAGAGAGAAAAAGAAAGAAATCTGAACAGTTAAAGNGT	2340
Db	2281	TTGTCAATCCAGCCTTCCAGAGAGAAAAAGAAAGAAATCTGAACAGTTAAAGNGT	2340
QY	2341	CTAATTAATGCTGAAGACCCCAAGATCTCATGTTAAAGTGGAAGGGTTTTGCAAACTG	2400
Db	2341	CTAATTAATGCTGAAGACCCCAAGATCTCATGTTAAAGTGGAAGGGTTTTGCAAACTG	2400
QY	2401	AAAGATCTGTAGAGATAGCAGTATTTTCATTGGTACTCGTAGTATTAATGGCACTAGC	2460
Db	2401	AAAGATCTGTAGAGATAGCAGTATTTTCATTGGTACTCGTAGTATTAATGGCACTAGC	2460
QY	2461	AAAGTATCTCGTACTCGGAAGTATAGCACTAGGAGGAGCCAAAACAGAACCAATTAAT	2520
Db	2461	AAAGTATCTCGTACTCGGAAGTATAGCACTAGGAGGAGCCAAAACAGAACCAATTAAT	2520
QY	2521	GTTGTAGTCACTGTGCAGCACTTTTGAAAAACCCCAAGGACTAATTCATGGTGTTCACAAAG	2580
Db	2521	GTTGTAGTCACTGTGCAGCACTTTTGAAAAACCCCAAGGACTAATTCATGGTGTTCACAAAG	2580
QY	2581	ATAATAGAAATGACACAGAAAGGCTTTAAAGTATCCATTGGGACATGAAGTTAACCACAGTC	2640
Db	2581	ATAATAGAAATGACACAGAAAGGCTTTAAAGTATCCATTGGGACATGAAGTTAACCACAGTC	2640
QY	2641	GGGAAACAGCATATGAATGGAAGAAAGTGAACCTGATGTCAGATATTTCAGAAATTCAT	2700
Db	2641	GGGAAACAGCATATGAATGGAAGAAAGTGAACCTGATGTCAGATATTTCAGAAATTCAT	2700
QY	2701	TCAAGGTTTCCAAAGCCGCGCAGTCATTTCTGCTGTTTCCAAATCCAGAAATGCGAAGAGG	2760
Db	2701	TCAAGGTTTCCAAAGCCGCGCAGTCATTTCTGCTGTTTCCAAATCCAGAAATGCGAAGAGG	2760
QY	2761	AATGTGCAACATTTCTGCCCACTCTCTGGGTCCTTAAAGAAACAAAGTCGCAAAAGTCACTT	2820
Db	2761	AATGTGCAACATTTCTGCCCACTCTCTGGGTCCTTAAAGAAACAAAGTCGCAAAAGTCACTT	2820
QY	2821	TTGAATGTGAACAAAGGAAAGAAATATAGGAAAGAAATGAGTCTAATATATAGCGCTAGC	2880
Db	2821	TTGAATGTGAACAAAGGAAAGAAATATAGGAAAGAAATGAGTCTAATATATAGCGCTAGC	2880
QY	2881	AGACAGTTAATATACATGACAGGCTTCTGCTGTGTTGGTCAGAAAGATTAAGCCAGTTGATA	2940
Db	2881	AGACAGTTAATATATACATGACAGGCTTCTGCTGTGTTGGTCAGAAAGATTAAGCCAGTTGATA	2940
QY	2941	ATGCCAAATGTAGTATCAAAAGAGGCTCTAGGTTTGTCTATCATCTCACTTTCAGAGGCA	3000
Db	2941	ATGCCAAATGTAGTATCAAAAGAGGCTCTAGGTTTGTCTATCATCTCACTTTCAGAGGCA	3000
QY	3001	ACGAAACTGGCAGCTATTAAGTCCAAATTAACATGAGACTTTTACAAAACCCATATCGTATAC	3060
Db	3001	ACGAAACTGGCAGCTATTAAGTCCAAATTAACATGAGACTTTTACAAAACCCATATCGTATAC	3060
QY	3061	CACACACTTTTCCCATTAAGTCAATTTGTAAACTAATTAAGAAAAATCTGCTAGAGG	3120
Db	3061	CACACACTTTTCCCATTAAGTCAATTTGTAAACTAATTAAGAAAAATCTGCTAGAGG	3120
QY	3121	AAAACTTTGAGGAACATTCATGTCACCTGAAGAGAAATGGGAATGAGAACATTTCCAA	3180

Db	3121	AAACCTTTGAGGAACATTCAATGTCACTCGAAGAGAAATGGCAATATGAAACATTTCCAA	3186
QY	3181	GTACAGTGAGCAAAATTAGCCGTAACTTAACATTTAGAGAAATGTTTTAAGAACGCACGT	3240
Db	3181	GTACAGTGAGCAAAATTAGCCGTAACTTAACATTTAGAGAAATGTTTTAAGAACGCACGT	3240
QY	3241	CAAGCAATATTAAATGAAGTAGTTCACAGTACTTAATGAAGTGCGCTCCAGTATTAAATGAA	3300
Db	3241	CAAGCAATATTAAATGAAGTAGTTCACAGTACTTAATGAAGTGCGCTCCAGTATTAAATGAA	3300
QY	3301	TAGGTTCCAGTGATATAAAACATTTCAAGCAGAACTAGGTGAAGAACAGAGAGGCCAAATTGA	3366
Db	3301	TAGGTTCCAGTGATATAAAACATTTCAAGCAGAACTAGGTGAAGAACAGAGAGGCCAAATTGA	3366
QY	3361	ATGCTATCTTAGATTAGAGGGGTTTTGCAACCTGAGGCTATTAAACAAAGTCTTCTGGAA	3426
Db	3361	ATGCTATCTTAGATTAGAGGGGTTTTGCAACCTGAGGCTATTAAACAAAGTCTTCTGGAA	3426
QY	3421	GTAATTGTAGCATCTCTGAATATAAAAGCAAGAAATATGAAGAAAGTAGTTCAGACTGTTA	3480
Db	3421	GTAATTGTAGCATCTCTGAATATAAAAGCAAGAAATATGAAGAAAGTAGTTCAGACTGTTA	3480
QY	3481	ATACGATTTTCTCTCATATCTGATTTAGATTACTTGAACAGCCTTTGGAGTAATC	3540
Db	3481	ATACGATTTTCTCTCATATCTGATTTAGATTACTTGAACAGCCTTTGGAGTAATC	3540
QY	3541	ATGCATCTCAGGTTGTTCTGTGAGACACCTGATGACCTTTAGATATGATGTGAATAAAG	3600
Db	3541	ATGCATCTCAGGTTGTTCTGTGAGACACCTGATGACCTTTAGATATGATGTGAATAAAG	3600
QY	3601	AAGATCTAGTTTTCCTGAATAATGACATTAAGAAAGTTTCGCTTTTATAGCAAAACCG	3666
Db	3601	AAGATCTAGTTTTCCTGAATAATGACATTAAGAAAGTTTCGCTTTTATAGCAAAACCG	3666
QY	3661	TCCAGAAAGGAGACTTACAGAGAGTCTAGCCCTTTCACCCATACACATTTGCTCAGG	3720
Db	3661	TCCAGAAAGGAGACTTACAGAGAGTCTAGCCCTTTCACCCATACACATTTGCTCAGG	3720
QY	3721	GTTACCGAAGAGGGGCCCAAGAAATTAGAGTCTCTCAAGAGACACTTAATCTGTAGAGATG	3780
Db	3721	GTTACCGAAGAGGGGCCCAAGAAATTAGAGTCTCTCAAGAGACACTTAATCTGTAGAGATG	3780
QY	3781	AAGAGCTTCCTGCTTCCACACACTTGTTATTTGTGAAGTAAACAAATTTACCTTCTCAGT	3840
Db	3781	AAGAGCTTCCTGCTTCCACACACTTGTTATTTGTGAAGTAAACAAATTTACCTTCTCAGT	3840
QY	3841	CTACTAGGCAATAGCACCGCTTCTACCGAGTCTGTCTTAAGAACACAGAGAGAAATTTAT	3900
Db	3841	CTACTAGGCAATAGCACCGCTTCTACCGAGTCTGTCTTAAGAACACAGAGAGAAATTTAT	3900
QY	3901	TATCTCTTAACAAATAGCTTAAATGATGCTGACGTAACCCAGTAAATTTGSCAAAGCATCTC	3966
Db	3901	TATCTCTTAACAAATAGCTTAAATGATGCTGACGTAACCCAGTAAATTTGSCAAAGCATCTC	3966
QY	3961	AGGAACATCACCTTAGTAGAGGAACAAATGTTCTGCTAGCTGTGTTTCTTCACAGTGCA	4020
Db	3961	AGGAACATCACCTTAGTAGAGGAACAAATGTTCTGCTAGCTGTGTTTCTTCACAGTGCA	4020
QY	4021	GTGAATTGGAAGACTTGACTGCAANTACAAACCCAGAGATCTTTCTTGAATGGTTCCT	4080
Db	4021	GTGAATTGGAAGACTTGACTGCAANTACAAACCCAGAGATCTTTCTTGAATGGTTCCT	4080
QY	4081	CCAAACAAATAGGCAATAGTCTGAACCCAGGAGGTTGTCGATAGTCTTCCACAGTGCA	4140
Db	4081	CCAAACAAATAGGCAATAGTCTGAACCCAGGAGGTTGTCGATAGTCTTCCACAGTGCA	4140
QY	4141	TTTCAGATGATGAAGAAAGAGAAACGGGCTTGAAGAAATTAATCAAGCAAGCAAAAGCA	4200
Db	4141	TTTCAGATGATGAAGAAAGAGAAACGGGCTTGAAGAAATTAATCAAGCAAGCAAAAGCA	4200
QY	4201	TGGATTCAAACTTAGTGAAGCAGCATCTGGGTGTGAAGTGAACAACCGCTCTCTCAG	4260
Db	4201	TGGATTCAAACTTAGTGAAGCAGCATCTGGGTGTGAAGTGAACAACCGCTCTCTCAG	4260





Db 241 ACATATTTTGCATAATTTTGCATGCTGGAACCTTCTCACACAGAGAAAGGCCCTTCACAGT 300  
OY 301 GTCCTTATGTAGATATATTAACCAAAAGAGCCTACAAAGAAAGTACGAGATTAGTC 360  
OY 301 GTCCTTATGTAGATATATTAACCAAAAGAGCCTACAAAGAAAGTACGAGATTAGTC 360  
Db 301 GTCCTTATGTAGATATATTAACCAAAAGAGCCTACAAAGAAAGTACGAGATTAGTC 360  
OY 361 AACTGTTGAAGAGCTATTGAAATCATTTGCTTTGACCTTGACACAGGTTTGGAGT 420  
Db 361 AACTGTTGAAGAGCTATTGAAATCATTTGCTTTGACCTTGACACAGGTTTGGAGT 420  
OY 421 ATGCAAAAGCTATATTTTGCATAAAGAAATTAATCTCTCCGTAACATCTAAAGATG 480  
Db 421 ATGCAAAAGCTATATTTTGCATAAAGAAATTAATCTCTCCGTAACATCTAAAGATG 480  
OY 481 AAGTTTCTATCATCCAAAGATATGGGCTACAGAACCCGTCGCAAAAAGACTTCTACGAGT 540  
Db 481 AAGTTTCTATCATCCAAAGATATGGGCTACAGAACCCGTCGCAAAAAGACTTCTACGAGT 540  
OY 541 AACCAGAAATCTCTCTGAGAGAAACCACTCAGTGTCCAACTCTCTAACTTGGAA 600  
Db 541 AACCAGAAATCTCTCTGAGAGAAACCACTCAGTGTCCAACTCTCTAACTTGGAA 600  
OY 601 CTGAGAGACTCTGAGAGCAAGAGCGGATACAACTCCAAAGAGCTCTCTACATG 660  
Db 601 CTGAGAGACTCTGAGAGCAAGAGCGGATACAACTCCAAAGAGCTCTCTACATG 660  
OY 661 AATTGGATCTGATTTCTCTGAGATACCGTTAATAGCAACTATTGCGAGTGGAG 720  
Db 661 AATTGGATCTGATTTCTCTGAGATACCGTTAATAGCAACTATTGCGAGTGGAG 720  
OY 721 ATCAAGAAATGTTCACAAATCAACCCCTCAAGAGAACCGAGGATGAATTCGATTG 780  
Db 721 ATCAAGAAATGTTCACAAATCAACCCCTCAAGAGAACCGAGGATGAATTCGATTG 780  
OY 781 CAAAAAGAGCTGTTGTAATTTTCTGAGAGAGATGTAACAATCTGTAACATCTAC 840  
Db 781 CAAAAAGAGCTGTTGTAATTTTCTGAGAGAGATGTAACAATCTGTAACATCTAC 840  
OY 841 CCAGTAATATGATTTGTAACACCACTGAGAGCGCTGAGAGCGATCCAGAAAGT 900  
Db 841 CCAGTAATATGATTTGTAACACCACTGAGAGCGCTGAGAGCGATCCAGAAAGT 900  
OY 901 ATCAGGAGTGTCTGTTTCAAACTGTCATGTGAGCCATGTGGCACAATATCTCATG 960  
Db 901 ATCAGGAGTGTCTGTTTCAAACTGTCATGTGAGCCATGTGGCACAATATCTCATG 960  
OY 961 GCTATTTACAGCAGAGAGACAGCACTTATCTACTCAATAGAGAGCAATGATGAGAA 1020  
Db 961 GCTATTTACAGCAGAGAGACAGCACTTATCTACTCAATAGAGAGCAATGATGAGAA 1020  
OY 1021 AGGCTGATTTCTGTAATTAAGCAAAAGCCTGCTTACAGAGAGCAATATGATGAG 1080  
Db 1021 AGGCTGATTTCTGTAATTAAGCAAAAGCCTGCTTACAGAGAGCAATATGATGAG 1080  
OY 1081 GGGCTGGAAGTAAAGCAATGTAATGATGAGCGGACTCCAGCAGAAAAAGGTAG 1140  
Db 1081 GGGCTGGAAGTAAAGCAATGTAATGATGAGCGGACTCCAGCAGAAAAAGGTAG 1140  
OY 1141 ATCTGAATGCTGATCCCTGTGTGAGAGAAAAAGATGGAATAGAGAAACTGCCATCT 1200  
Db 1141 ATCTGAATGCTGATCCCTGTGTGAGAGAAAAAGATGGAATAGAGAAACTGCCATCT 1200  
OY 1201 CAGAGAAATCTAGAGATCTAGAGATGTTCTTGATATACACTAATATGCGACATTTGCA 1260  
Db 1201 CAGAGAAATCTAGAGATCTAGAGATGTTCTTGATATACACTAATATGCGACATTTGCA 1260  
OY 1261 AAGTTAATGAGTGTTCAGAGAGTGAATGATGTTGATGATGATGATGATGATGATG 1320  
Db 1261 AAGTTAATGAGTGTTCAGAGAGTGAATGATGTTGATGATGATGATGATGATGATG 1320  
OY 1321 GGGAGTCTGATGCAAAAGCAAGTACGATGATGATGATGATGATGATGATGATGATG 1380  
Db 1321 GGGAGTCTGATGCAAAAGCAAGTACGATGATGATGATGATGATGATGATGATGATG 1380

OY 1381 AATATTTGTTCTTCAAGAGAAATAGACTTACTGGCCAGTGCATCTCATAGGCTTTAA 1440  
Db 1381 AATATTTGTTCTTCAAGAGAAATAGACTTACTGGCCAGTGCATCTCATAGGCTTTAA 1440  
OY 1441 TATGTAAGAGTGAAGAGTCTCAATATCTAGAGAGATATTTTGAAGCAAAATAT 1500  
Db 1441 TATGTAAGAGTGAAGAGTCTCAATATCTAGAGAGATATTTTGAAGCAAAATAT 1500  
OY 1501 TTGGGAAACCTATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560  
Db 1501 TTGGGAAACCTATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560  
OY 1561 TAAATATAGAGCAATTTGTTACTGAGCAGAGATTAATCAAGAGAGAGAGAGAGAG 1620  
Db 1561 TAAATATAGAGCAATTTGTTACTGAGCAGAGATTAATCAAGAGAGAGAGAGAGAG 1620  
OY 1621 AATTAAAGCGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680  
Db 1621 AATTAAAGCGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680  
OY 1681 CAGATTTGGAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740  
Db 1681 CAGATTTGGAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740  
OY 1741 AGAATGCTCAAGTCAATATTTACTAATAGTGCATAGAGATTAACCAAGAGTGAAT 1800  
Db 1741 AGAATGCTCAAGTCAATATTTACTAATAGTGCATAGAGATTAACCAAGAGTGAAT 1800  
OY 1801 CTATTCAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860  
Db 1801 CTATTCAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860  
OY 1861 AAACGAAGCTGAACCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920  
Db 1861 AAACGAAGCTGAACCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920  
OY 1921 ACAATTCAAAAGCAGCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980  
Db 1921 ACAATTCAAAAGCAGCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980  
OY 1981 ATGCGCTTGAAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 2040  
Db 1981 ATGCGCTTGAAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 2040  
OY 2041 TTGATAGTGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2100  
Db 2041 TTGATAGTGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2100  
OY 2101 GGCACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2160  
Db 2101 GGCACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2160  
OY 2161 GTACCAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220  
Db 2161 GTACCAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220  
OY 2221 ACTTAACAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2280  
Db 2221 ACTTAACAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2280  
OY 2281 TTGTCATCTCAAGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2340  
Db 2281 TTGTCATCTCAAGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2340  
OY 2341 CTAAATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2400  
Db 2341 CTAAATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2400  
OY 2401 AAAGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2460  
Db 2401 AAAGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2460

QY	2461	AAAGATCTCGTACTGGAAAGTTACCACTCTAAGGGAAGGCCAAAAAGCAACCAATTAAT	2520
Db	2461	AAAGATCTCGTACTGGAAAGTTACCACTCTAAGGGAAGGCCAAAAAGCAACCAATTAAT	2520
QY	2521	GTGTGAGTCAGTGTGACGATTTTGGAAAAAGGAGGAGCTAATTCATGTGTTTCCAAG	2580
Db	2521	GTGTGAGTCAGTGTGACGATTTTGGAAAAAGGAGGAGCTAATTCATGTGTTTCCAAG	2580
QY	2581	ATTAATAGAAATGACACACAGAAGGCTTTAAGTATCCATTGGGACATGAACTTAACCAAGTC	2640
Db	2581	ATTAATAGAAATGACACACAGAAGGCTTTAAGTATCCATTGGGACATGAACTTAACCAAGTC	2640
QY	2641	GGGAACACAGCATAGAAAAATGGAGAAAGTGAACCTGATGCTCAGATTTTGCAGAAATACAT	2700
Db	2641	GGGAACACAGCATAGAAAAATGGAGAAAGTGAACCTGATGCTCAGATTTTGCAGAAATACAT	2700
QY	2701	TCAGAGTTTCAAAAGCGCCAGTCATTTTCTCTGTTTTCAAATCCAGAAATGACAGAGAG	2760
Db	2701	TCAGAGTTTCAAAAGCGCCAGTCATTTTCTCTGTTTTCAAATCCAGAAATGACAGAGAG	2760
QY	2761	AATGTCGAACATTCCTGCGCCACCTCGGCTGCTTAAGAAACCAAGTCCAAAGTCACCT	2820
Db	2761	AATGTCGAACATTCCTGCGCCACCTCGGCTGCTTAAGAAACCAAGTCCAAAGTCACCT	2820
QY	2821	TTGAATGTGAACAAAAGGAGAAATATCAAGAAAGAAAGTGAAGTCTAATATCAAGCTGTAC	2880
Db	2821	TTGAATGTGAACAAAAGGAGAAATATCAAGAAAGAAAGTGAAGTCTAATATCAAGCTGTAC	2880
QY	2881	AGACAGTTAATATACCTGACGAGCTTCTCTGTTGGTGCAGAAAGATTAAGCACTGTATA	2940
Db	2881	AGACAGTTAATATACCTGACGAGCTTCTCTGTTGGTGCAGAAAGATTAAGCACTGTATA	2940
QY	2941	ATGCCAAATGTAGTATCAAAAGGAGGCTCTAGGTTTGTCTATCATCTCAGTTCCAGAGCA	3000
Db	2941	ATGCCAAATGTAGTATCAAAAGGAGGCTCTAGGTTTGTCTATCATCTCAGTTCCAGAGCA	3000
QY	3001	ACGAAATCGAGCTCATTTACTCCAAATTAACATGAGACTTTTTCAAAACCCATATGCTATAC	3060
Db	3001	ACGAAATCGAGCTCATTTACTCCAAATTAACATGAGACTTTTTCAAAACCCATATGCTATAC	3060
QY	3061	CACGACATTTTCCCATCAAGTCAATTTGTTAAATGTAATGTAAGAAAAATCTGCTAAGG	3120
Db	3061	CACGACATTTTCCCATCAAGTCAATTTGTTAAATGTAATGTAAGAAAAATCTGCTAAGG	3120
QY	3121	AAAACTTTGAGGAACATTCAATGTCCACCTGAAAGAAATGGGAATGAGAACATTTCCAA	3180
Db	3121	AAAACTTTGAGGAACATTCAATGTCCACCTGAAAGAAATGGGAATGAGAACATTTCCAA	3180
QY	3181	GTACAGTGAGACAAATTAAGCCGTAAATAACATTAGAGAAATGTTTAAAGAAAGCCAGCT	3240
Db	3181	GTACAGTGAGACAAATTAAGCCGTAAATAACATTAGAGAAATGTTTAAAGAAAGCCAGCT	3240
QY	3241	CAAGCAATTAATTAAGAGTAGTTCCAGTACATAATGAAGGCGCTCCAGTATTAATGA	3300
Db	3241	CAAGCAATTAATTAAGAGTAGTTCCAGTACATAATGAAGGCGCTCCAGTATTAATGA	3300
QY	3301	TAGGTTCCAGTGATGAAGAACTTCAACACAGAACTAGTATGAAGAGGCGCCAAATTTGA	3360
Db	3301	TAGGTTCCAGTGATGAAGAACTTCAACACAGAACTAGTATGAAGAGGCGCCAAATTTGA	3360
QY	3361	ATGCTATGCTTAGATTAGGGGTTTTTGGACCTGAGGCTCTAATAACAAGTCTTCTGGAA	3420
Db	3361	ATGCTATGCTTAGATTAGGGGTTTTTGGACCTGAGGCTCTAATAACAAGTCTTCTGGAA	3420
QY	3421	GTATTTGTAGCATCTCGAATTAATAAAGCAAGATATGAAGAATAGTTCCAGACTGTA	3480
Db	3421	GTATTTGTAGCATCTCGAATTAATAAAGCAAGATATGAAGAATAGTTCCAGACTGTA	3480
QY	3481	ATACAGATTTTCTCCATATCTGATTTTCAAGTAACTTGAAGAACCCATAGGAAAGTATGTC	3540
Db	3481	ATACAGATTTTCTCCATATCTGATTTTCAAGTAACTTGAAGAACCCATAGGAAAGTATGTC	3540
QY	3541	ATGCATCTCAGGTTTTGTTCTGAGACACCTGATGACCTGTTAGATGATGTTGAATAAAGG	3600

Db	3541	ATGCACTCAGGTTGGTCTGAGACACCTGATGACCTGTAGATGATGGTAATTAAGC	3600
QY	3601	AAGATACTAGTATTTGCTGAAATACATTAAGAAAGTTGCGTGTATTTAGCAAAAGC	3660
Db	3601	AAGATACTAGTATTTGCTGAAATATGACATTAAGAAAGTTGCGTGTATTTAGCAAAAGC	3660
QY	3661	TCCAAAGAGGAGAGCTTAGCAGAGAGTCTAGCCCTTCACCCATACATTTGGCTCAGG	3720
Db	3661	TCCAAAGAGGAGAGCTTAGCAGAGAGTCTAGCCCTTCACCCATACATTTGGCTCAGG	3720
QY	3721	GTTACCGAAGAGGGGCCAAGAAATTAGAGTCTAGAGAGAACTATCTAGGAGATG	3780
Db	3721	GTTACCGAAGAGGGGCCAAGAAATTAGAGTCTCTAGAGAGAACTATCTAGGAGATG	3780
QY	3781	AAGAGCTTCCCTGCTTCCACACTTGTATTGTGAAGTAACAATATACCTTCAGT	3840
Db	3781	AAGAGCTTCCCTGCTTCCACACTTGTATTGTGAAGTAACAATATACCTTCAGT	3840
QY	3841	CTACAGGCAATAGCAGCGTTGCTACGAGTGTCTGTAGAACAAGAGGGAATTTAT	3900
Db	3841	CTACAGGCAATAGCAGCGTTGCTACGAGTGTCTGTAGAACAAGAGGGAATTTAT	3900
QY	3901	TATCATTTGGAAGATAGCTTAAATGACTGACGTAAACAGGTAATATTGGCAAGGCAATC	3960
Db	3901	TATCATTTGGAAGATAGCTTAAATGACTGACGTAAACAGGTAATATTGGCAAGGCAATC	3960
QY	3961	AGGAACATACCTTAGTAGAGAAACAATAATGTTCTGTACCTGTTTCTTACAGTGA	4020
Db	3961	AGGAACATACCTTAGTAGAGAAACAATAATGTTCTGTACCTGTTTCTTACAGTGA	4020
QY	4021	GTTAATTTGGAAGACTTACTGCAATATCAAAACCAACCCAGATTCCTTCTGATTGTTCTT	4080
Db	4021	GTTAATTTGGAAGACTTACTGCAATATCAAAACCAACCCAGATTCCTTCTGATTGTTCTT	4080
QY	4081	CCAAACAATATAGGATCAGTCTTAAGCCAGGAGATTTGGTCTGATGACAAAGAAATGG	4140
Db	4081	CCAAACAATATAGGATCAGTCTTAAGCCAGGAGATTTGGTCTGATGACAAAGAAATGG	4140
QY	4141	TTTCAGATGATGAAGAAAGAGAAAGGCGCTTGGAGCAAAATATATCAAGAAAGCAAAACA	4200
Db	4141	TTTCAGATGATGAAGAAAGAGAAAGGCGCTTGGAGCAAAATATATCAAGAAAGCAAAACA	4200
QY	4201	TGGATTCAAATTTAGGTGAAGCAGCATCTGGGTGTAGAGTGAACAGAGCTCTCAAG	4260
Db	4201	TGGATTCAAATTTAGGTGAAGCAGCATCTGGGTGTAGAGTGAACAGAGCTCTCAAG	4260
QY	4261	ACTGTCAGGCGTATTCCTCAGAGTACATTTTAACCACTCAGCAGAGAGGATACCATGC	4320
Db	4261	ACTGTCAGGCGTATTCCTCAGAGTACATTTTAACCACTCAGCAGAGAGGATACCATGC	4320
QY	4321	AACATTAACCTGATTAAGCTCCAGCAGAGAAATGGCTAGACTAGAACCTGTGTTAGAACAC	4380
Db	4321	AACATTAACCTGATTAAGCTCCAGCAGAGAAATGGCTAGACTAGAACCTGTGTTAGAACAC	4380
QY	4381	ATGGGAGCCAGCCTTCTTAACAGCTACCCCTTCATATAGTGACITCTTGCCCTTGAAG	4440
Db	4381	ATGGGAGCCAGCCTTCTTAACAGCTACCCCTTCATATAGTGACITCTTGCCCTTGAAG	4440
QY	4441	ACCTGCAAAATCCGAACAAGCACATCAGAAAAACAGTATTAATCTCAGCAAAACATA	4500
Db	4441	ACCTGCAAAATCCGAACAAGCACATCAGAAAAACAGTATTAATCTCAGCAAAACATA	4500
QY	4501	GTGAATACCTTATATAGCCAGAAATCCAGAAAGCCTTCTCTCAGCAAGTTTGAAGGTCTG	4560
Db	4501	GTGAATACCTTATATAGCCAGAAATCCAGAAAGCCTTCTCTCAGCAAGTTTGAAGGTCTG	4560
QY	4561	CAGATAGTCTACAGATTAATAATAAGAACCAGAGTGGAAAGTATCCCTCTTAAT	4620
Db	4561	CAGATAGTCTACAGATTAATAATAAGAACCAGAGTGGAAAGTATCCCTCTTAAT	4620
QY	4621	GCCCATCATTAAGATGATAGTGTGATCATGCACAGTTGCTCTGGAGTCTTCAGATTAAGA	4680
Db	4621	GCCCATCATTAAGATGATAGTGTGATCATGCACAGTTGCTCTGGAGTCTTCAGATTAAGA	4680

Db 4621 GCCATCATATGATGATGAGTGTACATGACATGCTCTGAGAGCTTTCAGATAGAA 4680  
QY 4681 ACTACCATCTCTAAGAGAGAGCTCATTAAGTTGTTGATGTGAGAGACAACAGCTGGAG 4740  
Db 4681 ACTACCATCTCTAAGAGAGAGCTCATTAAGTTGTTGATGTGAGAGACAACAGCTGGAG 4740  
QY 4741 AGTCTGGGCGACACGATTTTGAGGGAACATCTTACTTGGCCAAAGCAAGATCTAGAGGAA 4800  
Db 4741 AGTCTGGGCGACACGATTTTGAGGGAACATCTTACTTGGCCAAAGCAAGATCTAGAGGAA 4800  
QY 4801 CCCCTTACTGGAATCTGCAATCAGCTCTCTCTATATGACCTGTAATCTGATCTCTG 4860  
Db 4801 CCCCTTACTGGAATCTGCAATCAGCTCTCTCTATATGACCTGTAATCTGATCTCTG 4860  
QY 4861 AAGACAGAGCCCGAGAGTACGCTCGTTGGCAACATACATCTTCAACCTCGCATTTGA 4920  
Db 4861 AAGACAGAGCCCGAGAGTACGCTCGTTGGCAACATACATCTTCAACCTCGCATTTGA 4920  
QY 4921 AAGTCCCAATGAAAGTTGCGAATCTGCCAGAGTCCAGCTGCTGCTCATACTACTG 4980  
Db 4921 AAGTCCCAATGAAAGTTGCGAATCTGCCAGAGTCCAGCTGCTGCTCATACTACTG 4980  
QY 4981 ATACTGCTGGTATATGCAATGGAAGAGTGTGACAGAGGAGAACCAAGATTTGACAG 5040  
Db 4981 ATACTGCTGGTATATGCAATGGAAGAGTGTGACAGAGGAGAACCAAGATTTGACAG 5040  
QY 5041 CTTCACAGAAAGGCTACACAAAGAAATGCTCATGCTGCTGCTGCTGCTGCTGCTGCTG 5100  
Db 5041 CTTCACAGAAAGGCTACACAAAGAAATGCTCATGCTGCTGCTGCTGCTGCTGCTGCTG 5100  
QY 5101 AATTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5160  
Db 5101 AATTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5160  
QY 5161 CTGAAAGAGTACTACTGTTGTTATGAAAAACAGATCTGAGTTGTTGTTGTTGTTGTTG 5220  
Db 5161 CTGAAAGAGTACTACTGTTGTTATGAAAAACAGATCTGAGTTGTTGTTGTTGTTGTTG 5220  
QY 5221 TGAATATTTTCTAGAAATTTGCGGAGAGAAATGAGTACTGTTGTTGTTGTTGTTGTTG 5280  
Db 5221 TGAATATTTTCTAGAAATTTGCGGAGAGAAATGAGTACTGTTGTTGTTGTTGTTGTTG 5280  
QY 5281 AGTCTATTAAGAAAGAAATGCTGATGAGCATGTTTGAAGTCAAGAGAGATGTGG 5340  
Db 5281 AGTCTATTAAGAAAGAAATGCTGATGAGCATGTTTGAAGTCAAGAGAGATGTGG 5340  
QY 5341 TCAATGAGAAACCCCAAGGTCCCAAGCAGACAGAGATCCCGAGACAGAAAGATCT 5400  
Db 5341 TCAATGAGAAACCCCAAGGTCCCAAGCAGACAGAGATCCCGAGACAGAAAGATCT 5400  
QY 5401 TCAGGGGCTAGAAATCTGCTGATGAGGCTTTCACCAATGCTCCAGATCACTGG 5460  
Db 5401 TCAGGGGCTAGAAATCTGCTGATGAGGCTTTCACCAATGCTCCAGATCACTGG 5460  
QY 5461 AATGATGATGAGTGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5520  
Db 5461 AATGATGATGAGTGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5520  
QY 5521 GCACAGGTGTGACCAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 5580  
Db 5521 GCACAGGTGTGACCAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 5580  
QY 5581 TCCATGCAATTTGGGAGATGTTGAGGACCTGTGTTGAGCCGAGAGTGGGTGTTGGAA 5640  
Db 5581 TCCATGCAATTTGGGAGATGTTGAGGACCTGTGTTGAGCCGAGAGTGGGTGTTGGAA 5640  
QY 5641 GTGTAGCACTCTACAGTCCAGAGAGTGTGACACCTGATATCCCAAGATCCCAAGATCC 5700  
Db 5641 GTGTAGCACTCTACAGTCCAGAGAGTGTGACACCTGATATCCCAAGATCCCAAGATCC 5700  
QY 5701 GGCACACTGGA 5711  
Db 5701 GGCACACTGGA 5711

RESULT 3  
AR033056  
LOCUS AR033056 5711 bp DNA linear PAT 29-SEP-1999  
DEFINITION Sequence 1 from patent US 5869245.  
ACCESSION AR033056  
VERSION AR033056.1 GI:5948661  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 5711)  
AUTHORS Yeung, A.T.  
TITLE Mismatch endonuclease and its use in identifying mutations in  
targeted polynucleotide strands  
JOURNAL Patent: US 5869245-A 1 09-FEB-1999;  
FEATURES  
source location/Qualifiers  
BASE COUNT 1956 a 1099 c 1274 g 1382 t  
ORIGIN  
Query Match 99.9%; Score 5707.8; DB 6; Length 5711;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 5709; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 AGCTGCTGAGACTTCTGAGACCCCGACAGGCTGTGGGTTTCTCAGATACCTGGCC 60  
Db 1 AGCTGCTGAGACTTCTGAGACCCCGACAGGCTGTGGGTTTCTCAGATACCTGGCC 60  
QY 61 CCGCGCTCAGAGAGGCTTACACCTCTGCTCTGAGGTAAGTTCATTGGAACAGAAAGAA 120  
Db 61 CCGCGCTCAGAGAGGCTTACACCTCTGCTCTGAGGTAAGTTCATTGGAACAGAAAGAA 120  
QY 121 TGGATTTATCTGCTCTTGGCGTTGAAGAAGTACAAATGCTATTATGCTATGACAGAAA 180  
Db 121 TGGATTTATCTGCTCTTGGCGTTGAAGAAGTACAAATGCTATTATGCTATGACAGAAA 180  
QY 181 TCTTAAAGTCCCATCTGCTGAGTGTATGATCAAGAACCTGCTCCCAAAAGTGTGACC 240  
Db 181 TCTTAAAGTCCCATCTGCTGAGTGTATCAAGAACCTGCTCCCAAAAGTGTGACC 240  
QY 241 ACATATTTTGAAGATTTTGCATGCTGAACCTTCTCAACAGAAAGAGGCTTCACTG 300  
Db 241 ACATATTTTGAAGATTTTGCATGCTGAACCTTCTCAACAGAAAGAGGCTTCACTG 300  
QY 301 GTCCCTTATGTAAGATGATATTAACCAAAAGAGCCTACAGAAAGTACAGATTTAGTC 360  
Db 301 GTCCCTTATGTAAGATGATATTAACCAAAAGAGCCTACAGAAAGTACAGATTTAGTC 360  
QY 361 AACTGCTGAAGAGATGTAAGAAATGCTGCTTTCAGTTGACAGAGTGTGGAGT 420  
Db 361 AACTGCTGAAGAGATGTAAGAAATGCTGCTTTCAGTTGACAGAGTGTGGAGT 420  
QY 421 ATGCAAAACAGTATATTTTGAAGAAAGAAATTAACCTCTCTGAACATCTTAAAGATG 480  
Db 421 ATGCAAAACAGTATATTTTGAAGAAAGAAATTAACCTCTCTGAACATCTTAAAGATG 480  
QY 481 AAGTTTCTATCTCAATCCAAATATGAGGCTACAGAAACCGTCCCAAAAGACTTTCAGAG 540  
Db 481 AAGTTTCTATCTCAATCCAAATATGAGGCTACAGAAACCGTCCCAAAAGACTTTCAGAG 540  
QY 541 AACCCGAATATCTTCTGAGAGAAACAGTCTGAGTCCAGCTCTAAGCTTAAAGTGA 600  
Db 541 AACCCGAATATCTTCTGAGAGAAACAGTCTGAGTCCAGCTCTAAGCTTAAAGTGA 600  
QY 601 CTGTGAGAACTCTGAGAGAAAGAGGAGTCAACCTTCAAAAGAGAGTGTCTACATG 660  
Db 601 CTGTGAGAACTCTGAGAGAAAGAGGAGTCAACCTTCAAAAGAGAGTGTCTACATG 660  
QY 661 AATTGGATCTGATTTCTTGAAGATACCGTTAATAGCAACTTATGCTGAGTGGAG 720  
Db 661 AATTGGATCTGATTTCTTGAAGATACCGTTAATAGCAACTTATGCTGAGTGGAG 720

Db 661 AATTGGATCTGATCTTCTGAGATACCGTTAATTAAGGCACTTATTCAGTGTGGAG 720  
QY 721 ATCAGAAATTGTTACAATACACCCCTCAGGAAACAGGATGAATCAGTTTGGATTCTG 780  
Db 721 ATCAGAAATTGTTACAATACACCCCTCAGGAAACAGGATGAATCAGTTTGGATTCTG 780  
QY 781 CAAAAAGGCTGCTTGTGAATTTTCTGAGACGATGTAACTGAATCTGAACATCATCAAC 840  
Db 781 CAAAAAGGCTGCTTGTGAATTTTCTGAGACGATGTAACTGAATCTGAACATCATCAAC 840  
QY 841 CCAGTAATATGATTTGAACACCCCTGAGAAAGCGCTGACGTGAGAGGATCCCAAAAAGT 900  
Db 841 CCAGTAATATGATTTGAACACCCCTGAGAAAGCGCTGACGTGAGAGGATCCCAAAAAGT 900  
QY 901 ATCAGGAGTACTGTTTCAAACTTGATGATGAGCCATGTCGACAAATCTCATGCA 960  
Db 901 ATCAGGAGTACTGTTTCAAACTTGATGATGAGCCATGTCGACAAATCTCATGCA 960  
QY 961 GCTCATTACAGCATGAGACAGCAGTTTATCTCACTAAAGACAGAAATGATAGAAA 1020  
Db 961 GCTCATTACAGCATGAGACAGCAGTTTATCTCACTAAAGACAGAAATGATAGAAA 1020  
QY 1021 AGGCTGAATCTGTAAATTAAGCAACAGCCTGGCTTAGAAGGACATCAAGAT 1080  
Db 1021 AGGCTGAATCTGTAAATTAAGCAACAGCCTGGCTTAGAAGGACATCAAGAT 1080  
QY 1081 GGGCTGGAAGTAAAGAAACATGTATGATAGGGGACTCCAGCACAGAAAAAGGTAG 1140  
Db 1081 GGGCTGGAAGTAAAGAAACATGTATGATAGGGGACTCCAGCACAGAAAAAGGTAG 1140  
QY 1141 ATCTGAATGTGATCCCTCTGTGAGAAAGAAAGATGAATAGCAAAAACCTGCATGCT 1200  
Db 1141 ATCTGAATGTGATCCCTCTGTGAGAAAGAAAGATGAATAGCAAAAACCTGCATGCT 1200  
QY 1201 CAGAGAAATCCTAGAACTACTGAGATGCTTGGATTAACATTAATAGAGCACTTGA 1260  
Db 1201 CAGAGAAATCCTAGAACTACTGAGATGCTTGGATTAACATTAATAGAGCACTTGA 1260  
QY 1261 AAGTTAATGATGTTTCCAGAAAGTATGATGATGATGATGATGATGATGATGATG 1320  
Db 1261 AAGTTAATGATGTTTCCAGAAAGTATGATGATGATGATGATGATGATGATGATG 1320  
QY 1321 GGGAGTGAATCAATGCCAAAGTACGTATGATGATGATGATGATGATGATGATGATG 1380  
Db 1321 GGGAGTGAATCAATGCCAAAGTACGTATGATGATGATGATGATGATGATGATGATG 1380  
QY 1381 AATATCTGCTTCTCAGAAATAGACTTACTGGCAGTATGATGATGATGATGATGATG 1440  
Db 1381 AATATCTGCTTCTCAGAAATAGACTTACTGGCAGTATGATGATGATGATGATGATG 1440  
QY 1441 TATGTAAAGTAAAGAGTTCACTCCAAATCAGTAGAGATTAATTTGAAGCAAAATAT 1500  
Db 1441 TATGTAAAGTAAAGAGTTCACTCCAAATCAGTAGAGATTAATTTGAAGCAAAATAT 1500  
QY 1501 TTGGGAAAACCTATCGGAAGAGGCAAGGCTCCCAACTTAAGCCATGTAACTGAATTC 1560  
Db 1501 TTGGGAAAACCTATCGGAAGAGGCAAGGCTCCCAACTTAAGCCATGTAACTGAATTC 1560  
QY 1561 TAATTATAGAGCATTTGTTACTGAGCCACAGATTAATACAGAGGCTCCCTCACAATA 1620  
Db 1561 TAATTATAGAGCATTTGTTACTGAGCCACAGATTAATACAGAGGCTCCCTCACAATA 1620  
QY 1621 AATTAAGGCTAAAGAGAGCTACATCAGGCTTCATCCTGAGATTTTATCAAAAAG 1680  
Db 1621 AATTAAGGCTAAAGAGAGCTACATCAGGCTTCATCCTGAGATTTTATCAAAAAG 1680  
QY 1681 CAGATTTGGCAGTTCAAAAAGACTCCTGAATGATTAATCAGGAACTAACCAAGGAGG 1740  
Db 1681 CAGATTTGGCAGTTCAAAAAGACTCCTGAATGATTAATCAGGAACTAACCAAGGAGG 1740  
QY 1741 AGAATGTCAGTGAATATTTACTAATAGTGTGATGATGATGATGATGATGATGATGAT 1800  
Db 1741 AGAATGTCAGTGAATATTTACTAATAGTGTGATGATGATGATGATGATGATGATGAT 1800

QY 1801 CTATTCAGAAATGAGAAAAATCCTAACCCAAATGATCAGTGAAGAAAAAGATCTGTTCA 1860  
Db 1801 CTATTCAGAAATGAGAAAAATCCTAACCCAAATGATCAGTGAAGAAAAAGATCTGTTCA 1860  
QY 1861 AAGGAAAGCTGAACCTTAATAGCAGAGTATTAAGCAATATGAACTGGAATTAATATCC 1920  
Db 1861 AAGGAAAGCTGAACCTTAATAGCAGAGTATTAAGCAATATGAACTGGAATTAATATCC 1920  
QY 1921 ACAATTTAAAGACCTTAATAAAGATAGGCTGAGAGGAAGTCTTACCAGCATATTC 1980  
Db 1921 ACAATTTAAAGACCTTAATAAAGATAGGCTGAGAGGAAGTCTTACCAGCATATTC 1980  
QY 1981 ATGGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040  
Db 1981 ATGGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040  
QY 2041 TTGATAGTTTCTAGCAGTGAAGAGATTAAGAAAAAGTACAACTGAATGCCAGTCA 2100  
Db 2041 TTGATAGTTTCTAGCAGTGAAGAGATTAAGAAAAAGTACAACTGAATGCCAGTCA 2100  
QY 2101 GGCACAGCAAAACCTTAACCTCATGAGAGTTAAAGACCTGCACTGAGCCAAAGAA 2160  
Db 2101 GGCACAGCAAAACCTTAACCTCATGAGAGTTAAAGACCTGCACTGAGCCAAAGAA 2160  
QY 2161 GTAAACAGCCAAATGAACAGACAGTAAAGACATGATGATGATGATGATGATGATGAT 2220  
Db 2161 GTAAACAGCCAAATGAACAGACAGTAAAGACATGATGATGATGATGATGATGATGAT 2220  
QY 2221 AGTTAAACAAATGACCTGTTCTTAAAGTGTCAATACAGTGAATTAAGAT 2280  
Db 2221 AGTTAAACAAATGACCTGTTCTTAAAGTGTCAATACAGTGAATTAAGAT 2280  
QY 2281 TTGTCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340  
Db 2281 TTGTCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340  
QY 2341 CTATTAATGCTGAAGACCCCAAGATCTCATGTTAAAGTGAAGAGGTTTGCAAACTG 2400  
Db 2341 CTATTAATGCTGAAGACCCCAAGATCTCATGTTAAAGTGAAGAGGTTTGCAAACTG 2400  
QY 2401 AAGATCTGATGAGAGTACAGATTAATGATGATGATGATGATGATGATGATGATGAT 2460  
Db 2401 AAGATCTGATGAGAGTACAGATTAATGATGATGATGATGATGATGATGATGATGAT 2460  
QY 2461 AAGATCTGATGAGAGTACAGATTAATGATGATGATGATGATGATGATGATGATGAT 2520  
Db 2461 AAGATCTGATGAGAGTACAGATTAATGATGATGATGATGATGATGATGATGATGAT 2520  
QY 2521 GTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2580  
Db 2521 GTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2580  
QY 2581 ATATATGAATGACACAGAAAGCTTAATGATGATGATGATGATGATGATGATGATGAT 2640  
Db 2581 ATATATGAATGACACAGAAAGCTTAATGATGATGATGATGATGATGATGATGATGAT 2640  
QY 2641 GGGAAACAGCATTAAGAAAGTGAAGAAAGTGAAGAAAGTGAAGAAAGTGAAGAAAG 2700  
Db 2641 GGGAAACAGCATTAAGAAAGTGAAGAAAGTGAAGAAAGTGAAGAAAGTGAAGAAAG 2700  
QY 2701 TCAAGTTTCAAGAGGCGCAGTATGCTGCTTTTCAAAATCCAGGAATGCAAGAGG 2760  
Db 2701 TCAAGTTTCAAGAGGCGCAGTATGCTGCTTTTCAAAATCCAGGAATGCAAGAGG 2760  
QY 2761 AATGTGCAACATTTCTGCCCAGTCTGGTCTTTAAAGAAACAAAGTCCAAAGTCACTT 2820  
Db 2761 AATGTGCAACATTTCTGCCCAGTCTGGTCTTTAAAGAAACAAAGTCCAAAGTCACTT 2820  
QY 2821 TTGAATGTGAACAAAGAGAAATTAAGAAAGATGATGATGATGATGATGATGATGAT 2880  
Db 2821 TTGAATGTGAACAAAGAGAAATTAAGAAAGATGATGATGATGATGATGATGATGAT 2880

OY 2881 AGACAGTAAATATACACTGACGGCTTTCTGTCGTGGTGTGTCAGAAAGATAGCCAGTTGATA 2940  
 |||||  
 Db 2881 AGACAGTAAATATACACTGACGGCTTTCTGTCGTGGTGTGTCAGAAAGATAGCCAGTTGATA 2940  
 OY 2941 ATGCCAAATGTATATCAAGAGAGGCTAGGTTTGTCTATCATCTCAGTTCCAGAGCA 3000  
 |||||  
 Db 2941 ATGCCAAATGTATATCAAGAGAGGCTAGGTTTGTCTATCATCTCAGTTCCAGAGCA 3000  
 OY 3001 AGCAACAGGAGCTCTTACTCCAAATTAACATGAGCTTTTACAAACCATATGCTATAC 3060  
 |||||  
 Db 3001 AGCAACAGGAGCTCTTACTCCAAATTAACATGAGCTTTTACAAACCATATGCTATAC 3060  
 OY 3061 CACCACCTTTTCCATCAAGTCATTTGTTAAACTAAATGTAGAAATATGCTAGAGG 3120  
 |||||  
 Db 3061 CACCACCTTTTCCATCAAGTCATTTGTTAAACTAAATGTAGAAATATGCTAGAGG 3120  
 OY 3121 AAACTTGGAGAACATTCATGTCACCTGAAAGAAATGGAAATGAGAACATTCCTCA 3180  
 |||||  
 Db 3121 AAACTTGGAGAACATTCATGTCACCTGAAAGAAATGGAAATGAGAACATTCCTCA 3180  
 OY 3181 GTACAGTAGACCAATTAGCCGTATATACATTTAGAGAAATGTTTTTAAAGAGCCAGCT 3240  
 |||||  
 Db 3181 GTACAGTAGACCAATTAGCCGTATATACATTTAGAGAAATGTTTTTAAAGAGCCAGCT 3240  
 OY 3241 CAAGCAATATTAATCAAGTAGGTTCCAGTACTAATGAAGTGGCTCCAGTATTAATGAA 3300  
 |||||  
 Db 3241 CAAGCAATATTAATCAAGTAGGTTCCAGTACTAATGAAGTGGCTCCAGTATTAATGAA 3300  
 OY 3301 TAGGTTCCAGTAGTAAACATTCACAGCAGAACTAGTAGAAACAGAGGCCAAATTTGA 3360  
 |||||  
 Db 3301 TAGGTTCCAGTAGTAAACATTCACAGCAGAACTAGTAGAAACAGAGGCCAAATTTGA 3360  
 OY 3361 ATGCTATGCTTAGATAGGGGTTTTGCAACCTGAGCTATTAACAAAGTCTCCTGSA 3420  
 |||||  
 Db 3361 ATGCTATGCTTAGATAGGGGTTTTGCAACCTGAGCTATTAACAAAGTCTCCTGSA 3420  
 OY 3421 GTAATTTAGCATCTCTGTAATTAAGCAAGATATGAAGAAAGTAGTTCAGACTGTTA 3480  
 |||||  
 Db 3421 GTAATTTAGCATCTCTGTAATTAAGCAAGATATGAAGAAAGTAGTTCAGACTGTTA 3480  
 OY 3481 ATACGATTTCTCCCATATCTGATTTAGATTAAGAACAGCTTATGGAAGTACTC 3540  
 |||||  
 Db 3481 ATACGATTTCTCCCATATCTGATTTAGATTAAGAACAGCTTATGGAAGTACTC 3540  
 OY 3541 ATGCTATCAGGTTTGTCTGAGACACCTGATGACCTGTTAGATGATGTTGAATTAAG 3600  
 |||||  
 Db 3541 ATGCTATCAGGTTTGTCTGAGACACCTGATGACCTGTTAGATGATGTTGAATTAAG 3600  
 OY 3601 AAGATACTAGTTTGTCTGAGAAATGACATTAAGGAAAGTTCTGCTGTTTTAGCAAAACG 3660  
 |||||  
 Db 3601 AAGATACTAGTTTGTCTGAGAAATGACATTAAGGAAAGTTCTGCTGTTTTAGCAAAACG 3660  
 OY 3661 TCCGAAAGAGAGCTTAGAGAGAGTCTTACGCCCTTACCCATACACATTTGGCTCAG 3720  
 |||||  
 Db 3661 TCCGAAAGAGAGCTTAGAGAGAGTCTTACGCCCTTACCCATACACATTTGGCTCAG 3720  
 OY 3721 GTTACCGAAGAGGGGCCAAGAAATTAGAGTCTCAGAGAGAACCTTATCTAGTGAAG 3780  
 |||||  
 Db 3721 GTTACCGAAGAGGGGCCAAGAAATTAGAGTCTCAGAGAGAACCTTATCTAGTGAAG 3780  
 OY 3781 AAGAGCTTCTGCTTCCACACTGTTATTTGTTAAAGTAACAAATTAACCTTCTCAGT 3840  
 |||||  
 Db 3781 AAGAGCTTCTGCTTCCACACTGTTATTTGTTAAAGTAACAAATTAACCTTCTCAGT 3840  
 OY 3841 CTACAGGATAGACACCTTGTCTACCGAGTCTCTCTAAGAACACAGAGAGAAATTTAT 3900  
 |||||  
 Db 3841 CTACAGGATAGACACCTTGTCTACCGAGTCTCTCTAAGAACACAGAGAGAAATTTAT 3900  
 OY 3901 TATCATTAAGAAATAGTTAATGATGACAGTAAACAGAGTATATTTGCAAGAGCATCTC 3960  
 |||||  
 Db 3901 TATCATTAAGAAATAGTTAATGATGACAGTAAACAGAGTATATTTGCAAGAGCATCTC 3960  
 OY 3961 AGGAACATCACCTTAGTGAGGAAACAAATATGTTCTGTAGCTGTTTTCTTACAGTGCA 4020  
 |||||

Db 3961 AGGAACATCACCTTAGTGAGGAAACAAATGTTGTGTAGCTTGTTTTCTTCCAGTGCA 4020  
 |||||  
 OY 4021 GTGAATTTGAGAGCTTTGACTGCAATTACAAACCCAGAGATCTTTTGTGATGTTCTT 4080  
 |||||  
 Db 4021 GTGAATTTGAGAGCTTTGACTGCAATTACAAACCCAGAGATCTTTTGTGATGTTCTT 4080  
 OY 4081 CCAAACAAATGAGGATCAGTGTGAAGCCAGAGGTTGTGTGATGTCAGAAAGTAAGG 4140  
 |||||  
 Db 4081 CCAAACAAATGAGGATCAGTGTGAAGCCAGAGGTTGTGTGATGTCAGAAAGTAAGG 4140  
 OY 4141 TTTGATGATGTAAGAAAGAGAACGGGCTTGGAAAGAAATATATCAGAGAGCAAGCA 4200  
 |||||  
 Db 4141 TTTGATGATGTAAGAAAGAGAACGGGCTTGGAAAGAAATATATCAGAGAGCAAGCA 4200  
 OY 4201 TGGATTTCAAACTTAGGTGTAAGAGCAGCATCTGGGTGTGAGAGTGAACAAAGCCCTCTGAAG 4260  
 |||||  
 Db 4201 TGGATTTCAAACTTAGGTGTAAGAGCAGCATCTGGGTGTGAGAGTGAACAAAGCCCTCTGAAG 4260  
 OY 4261 ACTGCTCAGGGCTATCTCTCTGAGAGTGAACATTTTAAACCTCAGCAGAGAGGATACATGC 4320  
 |||||  
 Db 4261 ACTGCTCAGGGCTATCTCTCTGAGAGTGAACATTTTAAACCTCAGCAGAGAGGATACATGC 4320  
 OY 4321 AACATTAACCTGATTAAGCTCCAGCAGAAATGGCTGAATAGAACCTGTGTAGAAGCAGC 4380  
 |||||  
 Db 4321 AACATTAACCTGATTAAGCTCCAGCAGAAATGGCTGAATAGAACCTGTGTAGAAGCAGC 4380  
 OY 4381 ATGGAGCCAGCTTCTTAACAGCTACCTTCCATCATTAAGTACGTCCTTGTGAGG 4440  
 |||||  
 Db 4381 ATGGAGCCAGCTTCTTAACAGCTACCTTCCATCATTAAGTACGTCCTTGTGAGG 4440  
 OY 4441 ACCTGGAAATCCAGCAAAAGCAGACATCAGAAAGAGCTATTAACCTCAGCAAGAACTA 4500  
 |||||  
 Db 4441 ACCTGGAAATCCAGCAAAAGCAGACATCAGAAAGAGCTATTAACCTCAGCAAGAACTA 4500  
 OY 4501 GTGAATTAACCTATAGCCAGAAATCCAGAGGCTTTCTGCTGACAAAGTTTGAAGTGTCTG 4560  
 |||||  
 Db 4501 GTGAATTAACCTATAGCCAGAAATCCAGAGGCTTTCTGCTGACAAAGTTTGAAGTGTCTG 4560  
 OY 4561 CAGATAGTTCTACCGTAAATTAAGAACCCAGAGGTGGAAGGTCATCCCTTGAAT 4620  
 |||||  
 Db 4561 CAGATAGTTCTACCGTAAATTAAGAACCCAGAGGTGGAAGGTCATCCCTTGAAT 4620  
 OY 4621 GCCCATATTAAGATATAGTGTGATGATGACAGAGTGTCTGAGAGCTTTCAGATTAAG 4680  
 |||||  
 Db 4621 GCCCATATTAAGATATAGTGTGATGATGACAGAGTGTCTGAGAGCTTTCAGATTAAG 4680  
 OY 4681 ACTACCCATCTCAGAGAGGCTCATTTAAGGTTGTGATGTGAGAGGCAACAGCTGGAAG 4740  
 |||||  
 Db 4681 ACTACCCATCTCAGAGAGGCTCATTTAAGGTTGTGATGTGAGAGGCAACAGCTGGAAG 4740  
 OY 4741 AGTCTGGGCCACAGCATTTTGAAGGAAACATCTTACTTCCAAAGGCAAAATCTAGAGGAA 4800  
 |||||  
 Db 4741 AGTCTGGGCCACAGCATTTTGAAGGAAACATCTTACTTCCAAAGGCAAAATCTAGAGGAA 4800  
 OY 4801 CCCCTTACCTGGAATCTGGAATCAGCTCTCTCTGATGAGACCCCTGAATCTGCTCTG 4860  
 |||||  
 Db 4801 CCCCTTACCTGGAATCTGGAATCAGCTCTCTCTGATGAGACCCCTGAATCTGCTCTG 4860  
 OY 4861 AAGACAGAGCCCAAGAGTCAAGCTCTGTTGGCAACATACATCTTCAACCTCTGCAATGA 4920  
 |||||  
 Db 4861 AAGACAGAGCCCAAGAGTCAAGCTCTGTTGGCAACATACATCTTCAACCTCTGCAATGA 4920  
 OY 4921 AAGTTTCCCAATTTGAAGTTTGCAGAAATCTGCCAAGTCCAGCTGCTCATATCTACTG 4980  
 |||||  
 Db 4921 AAGTTTCCCAATTTGAAGTTTGCAGAAATCTGCCAAGTCCAGCTGCTCATATCTACTG 4980  
 OY 4981 ATACTGCTGGGTATTAATCAATGAGAAAGTGTGAGAGGAGAGAGGCAAAATTTGACAG 5040  
 |||||  
 Db 4981 ATACTGCTGGGTATTAATCAATGAGAAAGTGTGAGAGGAGAGAGGCAAAATTTGACAG 5040  
 OY 5041 CTTCAACGAAGAGGTCAACAAAGAAATGTCCATGTGTGTCTGAGCTGAGCCCAAG 5100  
 |||||

```

Db      5041 CTTCACAGAAAGGCTCACAAAAGAAATGTCATGATGCTGCTGCGCTGACCCAGAG 5100
OY      5101 AATTTATGCTGCTAGCAAGTTTGCAGAAAACACACATCTTAATCTAATCTAATTA 5160
Db      5101 AATTTATGCTGCTAGCAAGTTTGCAGAAAACACACATCTTAATCTAATCTAATTA 5160
OY      5161 CTGAGAGACTACTCATGTTGTTATGAAAACAGATGCTGAGTTGTGTGAGAGGACAC 5220
Db      5161 CTGAGAGACTACTCATGTTGTTATGAAAACAGATGCTGAGTTGTGTGAGAGGACAC 5220
OY      5221 TGAATATTTTCTAGCAATTCGGGAGAGAAAATGGTAGTACATTTCTGGGTGACC 5280
Db      5221 TGAATATTTTCTAGCAATTCGGGAGAGAAAATGGTAGTACATTTCTGGGTGACC 5280
OY      5281 AGCTATTTAAAGAAAGAAAATGCTGAATGAGATGATTTTGAAGTCCAGAGAGATGAG 5340
Db      5281 AGCTATTTAAAGAAAGAAAATGCTGAATGAGATGATTTTGAAGTCCAGAGAGATGAG 5340
OY      5341 TCAATGGAAGAAACACCAAGGTCACAAAGCAGACAGAGAATCCAGAGACAGAAAGATCT 5400
Db      5341 TCAATGGAAGAAACACCAAGGTCACAAAGCAGACAGAGAATCCAGAGACAGAAAGATCT 5400
OY      5401 TCAGGGGGCTAGAAATCTGTTGCTATGGCCCTTACCACATGCCCCAGATCACTGG 5460
Db      5401 TCAGGGGGCTAGAAATCTGTTGCTATGGCCCTTACCACATGCCCCAGATCACTGG 5460
OY      5461 AATGATGATGACAGCTGTGCTGCTCTGATGAGAGAGAGCTTCATCATTCACCCCTTG 5520
Db      5461 AATGATGATGACAGCTGTGCTGCTCTGATGAGAGAGAGCTTCATCATTCACCCCTTG 5520
OY      5521 GCACAGGTGTCAACCCCAATTTGTGTTGTGACAGCAGATGCTTGACAGAGACAAATGGCT 5580
Db      5521 GCACAGGTGTCAACCCCAATTTGTGTTGTGACAGCAGATGCTTGACAGAGACAAATGGCT 5580
OY      5581 TCCATGCAATTTGGGCAATGTGTGAGCAGCTGTGTGAGACCCGAGAGTGGTGTGGACA 5640
Db      5581 TCCATGCAATTTGGGCAATGTGTGAGCAGCTGTGTGAGACCCGAGAGTGGTGTGGACA 5640
OY      5641 GTGTAGCACTCTACAGTGCAGAGAGAGTGGACACCTACCTGATACCCAGATGCCCCACA 5700
Db      5641 GTGTAGCACTCTACAGTGCAGAGAGAGTGGACACCTACCTGATACCCAGATGCCCCACA 5700
OY      5701 GCCCACTACTGA 5711
Db      5701 GCCCACTACTGA 5711
RESULT 4
HSU14680 5711 bp mRNA linear PRI 10-JUN-2002
LOCUS Homo sapiens breast and ovarian cancer susceptibility (BRCA1) mRNA,
DEFINITION complete cds.
ACCESSION U14680.1 GI:555931
VERSION 1
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 5711)
AUTHORS Mikl,Y., Swensen,J., Shattuck-Eidens,D., Futreal,P.A., Harshman,K.,
Tavtigian,S., Liu,Q., Cochran,C., Bennett,L.M., Ding,W., Bell,R.,
Rosenthal,J., Hussey,C., Tran,T., McClure,M., Frye,C., Hattier,T.,
Phelps,R., Haugen-Strano,A., Katcher,H., Yakumo,K., Gholami,Z.,
Shaffer,D., Stone,S., Bayer,S., Wray,C., Bogden,R., Dayananth,P.,
Ward,J., Tonin,P., Narod,S., Bristow,P.K., Norris,F.H.,
Heweling,L., Morrison,P., Kosteck,P., Lai,M., Barrettl,J.C.,
Lewis,C., Neuhausen,S., Cannon-Albright,L., Goldgar,D., Wiseman,R.,
Kamb,A. and Skolnick,M.H.
TITLE A strong candidate for the breast and ovarian cancer susceptibility
gene BRCA1
JOURNAL Science 266 (5182), 66-71 (1994)
MEDLINE 95023896

```

```

PUBMED 7545954
REFERENCE 2 (bases 1 to 5711)
AUTHORS Skolnick,M.H.
TITLE Direct Submission
JOURNAL Submitted (14-SEP-1994) Mark H. Skolnick, Myriad Genetics Inc. and
the University of Utah, 421 Wakara Way, Suite 201, Salt Lake City,
UT 84108, USA
FEATURES
source Location/Qualifiers
1..5711
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17q21; spans D17S855"
/note="For sequence of alternatively spliced exon 4, see
GenBank Accession Number U15955"
1..5711
/gene="BRCA1"
1..119
/gene="BRCA1"
1..100
/gene="BRCA1"
/number=1
101..199
/gene="BRCA1"
/number=2
120..5711
/gene="BRCA1"
/note="Influences susceptibility to breast and ovarian
cancer."
/codon_start=1
/product="breast and ovarian cancer susceptibility"
/protein_id="AA473985.1"
/db_xref="GI:555932"
/translation="MDLSALRFEVNVNINAMKILIECPLELIEKPEVSTCDHIFC
KFCMLKLNKKQPSQCPCKNDITRSLDSETRSPQVLEBLKIIICAFQDITGLEYA
NSYNFAKRNNSPEHLKDEYSIIQSMGVRRARARLQSEPEPMSLOETSLVQLNLSG
TVRLIRTKRQIQPKTSVYIELSDSEEDVKNATVCSYGDDELLQITPQGRDEISL
DSAKKACESEETDVNTTEHHQPSNNDLNTTERAARERHPKVGSSVNLVPECGT
NTHASSLOHENSLLLTQKDRMNERKAFKCSKQPLARSOHNRNAGSECTNDNRP
STERKYDNLNAPLCEERENKOKLPCSEMDPRDEVPYTLTNSLOKVNEMPSRDEL
LGSDDSHDGSSESNKAVADYDVLNEVDYSGSEKIDILASDPHALCKSERHRSK
SVESNIEDKLEGTYYRKASLPVLSHVENLIIGAVTEPQIIQERPLTKLRRRP
TSGLHPEDFLIKADLAVOKTPPEMNOGTQEDONGVNNITNSGHENKTKSLRER
NPNPIEIEKESAFTKAEPISSISNMELEINHNSKAPKRNLRKRSSTHIALE
LVYSRNLSPPNCTEIOIDSSSEELIKKKKKYNOMPVHRNRLQMGKRPATGAKKN
KPNQOTSKRSDPTPEPKLTNAPGSPCTSNSELEKFEVNSLPREEKELEVKY
SNNAEDKDLMSGERVLOTENSVSSSISLVPGDTGTQESTISLEVSTLKAETP
NKCVSQCAAFENPKGLIHGCSKDNNDTEGFKYPLIGHVNSHRETSIEEESLDAQY
LQNTFKVSKRQSFAPSPNPGNAEECATFSAHGSLKQSPKVTPECEDEKESNOCKE
SNIRPVQVNTITAGFVYQKDKPVDNACKSIGSGRPELSQFRNENGLITPKHG
LQNPVYRIPPLPIFKSPVTKCKKLLLENFEHSHSPREMGNNIPIVSTISRN
IRENVFEASNSNINFGVSGSTNEVSSINEDIGSDENIOAEIQRNGPPLNMLFLGY
LOPEVYKQSLPGSNCKHPEIKKQIEVEVQVYVTDPSPIILSDNLEOPMGSSHAQVC
SETPDDLLDDELKEDTSEFANDIKESAFVSKSVQGLSLRSPFTTHLHAQYGR
GAKLSESEENLSSDEBELPCFQHLLEGVANNIPSSQSTSHRYATELCKRNEENLIS
LKNLSLNCNSQVILAKASQEHLSSETKCSASFSSQCESELDTANTTOPPLIGS
SKOMRHOSESQVGLSDKELVSDPERGGLFENNEDOSMSNLCGEASGCESTSV
SPQCSGISQSDSLITLQOQRDPTQHNLIKIQOEVALEAEVLQDHGSPNSVYSIIDS
SALEDLRNPQOSTSEKAVLISQKSEYPISONPEGISAKFEVSDSSSKRKEGVE
RSSPSKQPSLDDRWYHNSGSLQNRNYSQELKLVNIVEQDQLESBPDLTLEST
LPQDLEGTPYLESGISLFSDESDPSDBRAESRANVIPSISALKVPKLAQES
AQPAAHATDTDGTAGYNAMEBSYREKPELTASTERYNKMMSVYSGLVPEEMLYKFR
ARKHNTLTLTLEETHVVMKDAEFVCEERTIKYELGJAGKVVYVYVWYSIKER
KLNENHDFEVRGQVYVGRNHQCPKRARESODRKIPGLETICCYGPTNMPDQLEMY
QLCGASVKEHLSFTLGTGVHPLVVYQPDAMTEDNGFAHIGQMCAPVYTRKVLDSV
ALVQCEDELDTYLLPQIPHSY"
200..253
/gene="BRCA1"
/number=3
254..331
/gene="BRCA1"
/number=5
332..420
exon
exon
exon

```



```

/ gene="BRCA1"
/ number=6
421. .560
/ gene="BRCA1"
/ number=7
561. .665
/ gene="BRCA1"
/ number=8
666. .712
/ gene="BRCA1"
/ number=9
713. .788
/ gene="BRCA1"
/ number=10
789. .4215
/ gene="BRCA1"
/ number=11
4216. .4302
/ gene="BRCA1"
/ number=12
4303. .4476
/ gene="BRCA1"
/ number=13
4477. .4603
/ gene="BRCA1"
/ number=14
4604. .4794
/ gene="BRCA1"
/ number=15
4795. .5105
/ gene="BRCA1"
/ number=16
5106. .5193
/ gene="BRCA1"
/ number=17
5194. .5273
/ gene="BRCA1"
/ number=18
5274. .5310
/ gene="BRCA1"
/ number=19
5311. .5396
/ gene="BRCA1"
/ number=20
5397. .5451
/ gene="BRCA1"
/ number=21
5452. .5526
/ gene="BRCA1"
/ number=22
5527. .5586
/ gene="BRCA1"
/ number=23
5587. .5711
/ gene="BRCA1"
/ number=24
1956 a 1099 c 1274 g 1382 t
BASE COUNT
ORIGIN
Query Match 99.9%; Score 5707.8; DB 9; Length 5711;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5709; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

121 TGGATTATCTGCTCTGCGGTGGAAGAGTACAAATGTCAATATGCTATGCAAGAAA 180
181 TCTTAGAGTCCCACTGCTGAGTGTATCAAGAACCTGTCTCCCAAGAGTGACC 240
181 TCTTAGAGTCTGCTGCTGAGTGTATCAAGAACCTGTCTCCCAAGAGTGACC 240
241 ACATATTTTGCAGAAATTTTGCATGCTGAACCTCTCCACAGAGAGAGGCTTCACAGT 300
241 ACATATTTTGCAGAAATTTTGCATGCTGAACCTCTCCACAGAGAGAGGCTTCACAGT 300
241 ACATATTTTGCAGAAATTTTGCATGCTGAACCTCTCCACAGAGAGAGGCTTCACAGT 300
301 GTCCCTTTATGTATAGATATATACCAAAAGAGGCTTCAAGAAAGTACAGATTAGTC 360
301 GTCCCTTTATGTATAGATATATACCAAAAGAGGCTTCAAGAAAGTACAGATTAGTC 360
361 AACTGTGTAAGAGCTATTTGAAAATCTTGTGCTTTCAGCTTACACAGAGTTGGAGT 420
361 AACTGTGTAAGAGCTATTTGAAAATCTTGTGCTTTCAGCTTACACAGAGTTGGAGT 420
421 ATGCAGAGCTATATATTTTGCAGAAATTTTGCATGCTGAACCTTCAAGAGATG 480
421 ATGCAGAGCTATATATTTTGCAGAAATTTTGCATGCTGAACCTTCAAGAGATG 480
481 AAGTTTCTATATCCAAAGTATGGCTACAGAAACCGTGCCAAAGACTTCTACAGATG 540
481 AAGTTTCTATATCCAAAGTATGGCTACAGAAACCGTGCCAAAGACTTCTACAGATG 540
541 AACCCGAAATCCCTTCTGAGAGAAACCACTTCTACAGTCTTCAACCTTGGAA 600
541 AACCCGAAATCCCTTCTGAGAGAAACCACTTCTACAGTCTTCAACCTTGGAA 600
601 CTGTGAGACTCTGAGAGCAAAAGCAGGATATCAACCTCAAAAGAGCTCTTACATTTG 660
601 CTGTGAGAGACTCTGAGAGCAAAAGCAGGATATCAACCTCAAAAGAGCTCTTACATTTG 660
661 AATTGGATCTGATTTCTCTAGAGTACCGTTAATAGGCACCTTATTCAGTGTGGAG 720
661 AATTGGATCTGATTTCTCTAGAGTACCGTTAATAGGCACCTTATTCAGTGTGGAG 720
721 ATCAAGATTTGTTCAATCAACCCCTCAAGAGACCGAGATGAAATCGATTGGATTCTG 780
721 ATCAAGATTTGTTCAATCAACCCCTCAAGAGACCGAGATGAAATCGATTGGATTCTG 780
781 CAAAAAGGCTGCTTGTGAATTTTCTGAGAGGATGTATCAAAATACATGATCAAC 840
781 CAAAAAGGCTGCTTGTGAATTTTCTGAGAGGATGTATCAAAATACATGATCAAC 840
841 CCAGTAATATGATTTTGAACACCACTGAGAGCGTGAGAGGATCCAGAAAGT 900
841 CCAGTAATATGATTTTGAACACCACTGAGAGCGTGAGAGGATCCAGAAAGT 900
901 ATCAGGATGTTCTGTTCAACCTTGCATGTGAGAGCCATGTGCAAAATCTCATGCCA 960
901 ATCAGGATGTTCTGTTCAACCTTGCATGTGAGAGCCATGTGCAAAATCTCATGCCA 960
961 GCTCATTACAGCAGAGAACGAGCTTATTAATACCAAGATGAAATGATGATGAGAAA 1020
961 GCTCATTACAGCAGAGAACGAGCTTATTAATACCAAGATGAAATGATGATGAGAAA 1020
1021 AGGCTGAATTTCTGTAATTAAGCAACAGCCTGCTTACAGAGAGCCACATCAACAT 1080
1021 AGGCTGAATTTCTGTAATTAAGCAACAGCCTGCTTACAGAGAGCCACATCAACAT 1080
1081 GGGCTGGAAGTAAGAAACATGATATAGTAGCGAGCTCCAGACAGAAAAAAGGATG 1140
1081 GGGCTGGAAGTAAGAAACATGATATAGTAGCGAGCTCCAGACAGAAAAAAGGATG 1140
1141 ATCTGAATGCTGATCCCTGCTGTGAGAGAAAGATGAAATGATGATGATGATGATG 1200
1141 ATCTGAATGCTGATCCCTGCTGTGAGAGAAAGATGAAATGATGATGATGATGATG 1200
1201 CAGAGAAATCTTGAAGATAGTGAAGATGCTTCTGAGATTAACCTAATAGCAGATTGCA 1260
1201 CAGAGAAATCTTGAAGATAGTGAAGATGCTTCTGAGATTAACCTAATAGCAGATTGCA 1260

```

QY 1261 AACGTAAATGAGTGGTTTTCCAGAAAGTATGTAAGTGTTCGATGACTCAGATGATG 1320  
|||||  
Db 1261 AAGTTATATGAGTGGTTTTCCAGAAAGTATGTAAGTGTTCGATGACTCAGATGATG 1320  
QY 1321 GGGAGTCTGATCAATAATGCCAAAGTACGTATGTTGGACGTTCTAATAGAGTATG 1380  
|||||  
Db 1321 GGGAGTCTGATCAATAATGCCAAAGTACGTATGTTGGACGTTCTAATAGAGTATG 1380  
QY 1381 AATATTCTGTTCTTCGAGAAAATAGACTTACGGCAGATGAGCTCATGAGGCTTTAA 1440  
|||||  
Db 1381 AATATTCTGTTCTTCGAGAAAATAGACTTACGGCAGATGAGCTCATGAGGCTTTAA 1440  
QY 1441 TATGTAAAGTGAAGAAGTTCACCTCAATCAGTAGAGAGTAATATTTGAACACAAATAT 1500  
|||||  
Db 1441 TATGTAAAGTGAAGAAGTTCACCTCAATCAGTAGAGAGTAATATTTGAACACAAATAT 1500  
QY 1501 TTGGGAAAACCTATCGGAAAGAGGCAAGCCTCCCAACTTAAGCCATGTAACTGAAAATC 1560  
|||||  
Db 1501 TTGGGAAAACCTATCGGAAAGAGGCAAGCCTCCCAACTTAAGCCATGTAACTGAAAATC 1560  
QY 1561 TAATTTATGAGAGATTTGTTACTGAGCCACAGATATACAAAGAGCGTCCCTCACAATA 1620  
|||||  
Db 1561 TAATTTATGAGAGATTTGTTACTGAGCCACAGATATACAAAGAGCGTCCCTCACAATA 1620  
QY 1621 AATTAAAGCTTAAAGAGAGACTACATCAGGCTTCATCCGAGAGATTTTATCAAGAAAG 1680  
|||||  
Db 1621 AATTAAAGCTTAAAGAGAGACTACATCAGGCTTCATCCGAGAGATTTTATCAAGAAAG 1680  
QY 1681 CAGATTGGCAGTTCAAAAGAGCTCTGTAATATGATAAATCAGGAGCTAACCAAGAGAGC 1740  
|||||  
Db 1681 CAGATTGGCAGTTCAAAAGAGCTCTGTAATATGATAAATCAGGAGCTAACCAAGAGAGC 1740  
QY 1741 AGAATGTCAGATGATGAATATTAATAGTGTGTCAGTAATTAACAAAGAGTAT 1800  
|||||  
Db 1741 AGAATGTCAGATGATGAATATTAATAGTGTGTCAGTAATTAACAAAGAGTAT 1800  
QY 1801 CTATTCAGATGAGAAAATCTTAACCCATAGATCCTCGAAAAGAAATCTGCTTTCA 1860  
|||||  
Db 1801 CTATTCAGATGAGAAAATCTTAACCCATAGATCCTCGAAAAGAAATCTGCTTTCA 1860  
QY 1861 AAACGAAAGCTGAACCTATTAAGCAGAGTATAGCAATATGAACTGGAATTTAAATTC 1920  
|||||  
Db 1861 AAACGAAAGCTGAACCTATTAAGCAGAGTATAGCAATATGGAATTTAAATTC 1920  
QY 1921 ACAATTCAAAAGACCTTAAAGAAATAGGCTGAGGAGAGTCTTCTACAGGCAATTC 1980  
|||||  
Db 1921 ACAATTCAAAAGACCTTAAAGAAATAGGCTGAGGAGAGTCTTCTACAGGCAATTC 1980  
QY 1981 ATGCGCTTGAAGTATAGTCAATCTAAGCCACCTAATTTGACTGAAATTCGAAA 2040  
|||||  
Db 1981 ATGCGCTTGAAGTATAGTCAATCTAAGCCACCTAATTTGACTGAAATTCGAAA 2040  
QY 2041 TTGATATGTTGTTTACAGAGTGAAGAGATTAAGAAAAAAAGTACAAACCAATGCGACATCA 2100  
|||||  
Db 2041 TTGATATGTTGTTTACAGAGTGAAGAGATTAAGAAAAAAAGTACAAACCAATGCGACATCA 2100  
QY 2101 GGCACAGCAGAAACCTACACACTCATGGAAGGTAAAGAACTGGAAGTGGAGCCAGAGAGA 2160  
|||||  
Db 2101 GGCACAGCAGAAACCTACACACTCATGGAAGGTAAAGAACTGGAAGTGGAGCCAGAGAGA 2160  
QY 2161 GTAACAGGCCAATGAACAGACAAATGAAGAGATGACAGTACTTCTCCAGAGCTGA 2220  
|||||  
Db 2161 GTAACAGGCCAATGAACAGACAAATGAAGAGATGACAGTACTTCTCCAGAGCTGA 2220  
QY 2221 AGTTAAACAATGACAGCTGGTCTTTTACTAAGTTCCTCAATACAGTAAGCAATTAAGCAAT 2280  
|||||  
Db 2221 AGTTAAACAATGACAGCTGGTCTTTTACTAAGTTCCTCAATACAGTAAGCAATTAAGCAAT 2280  
QY 2281 TTGTCAATCTAGCTTTCCAAAGAGAGAAAAAGAGAACTAGAAACAGTTTAAAGTGT 2340  
|||||  
Db 2281 TTGTCAATCTAGCTTTCCAAAGAGAGAAAAAGAGAACTAGAAACAGTTTAAAGTGT 2340

QY 2341 CTAATAATGCTGAAGACCCCAAGATCTCATGTTAACTGAGAGAAAGGTTTTGCAAACTG 2400  
|||||  
Db 2341 CTAATAATGCTGAAGACCCCAAGATCTCATGTTAACTGAGAGAAAGGTTTTGCAAACTG 2400  
QY 2401 AAAGATCTGTAAGAGTACAGATTTTCAATTTGCTGACTGGTACTGATTTATGCACTCAG 2460  
|||||  
Db 2401 AAAGATCTGTAAGAGTACAGATTTTCAATTTGCTGACTGGTACTGATTTATGCACTCAG 2460  
QY 2461 AAAGATCTGTTACTGGAAGTTAGCACCTGAGGAGAAAGCAAAACAGAACTAAAT 2520  
|||||  
Db 2461 AAAGATCTGTTACTGGAAGTTAGCACCTGAGGAGAAAGCAAAACAGAACTAAAT 2520  
QY 2521 GTGTGACTGAGTGCAGATTTGAAAAACCCCAAGGACTAATTCATGTTGTTCCAAAG 2580  
|||||  
Db 2521 GTGTGACTGAGTGCAGATTTGAAAAACCCCAAGGACTAATTCATGTTGTTCCAAAG 2580  
QY 2581 ATATATGAATATGACACAGAAAGGCTTTAATGATTCATTGGGACATGAAGTTAACACAGTC 2640  
|||||  
Db 2581 ATATATGAATATGACACAGAAAGGCTTTAATGATTCATTGGGACATGAAGTTAACACAGTC 2640  
QY 2641 GGGAAACAGCATGGAATGGAAGAAAGTGAACCTGATGCTGATTTGCAAGATACAT 2700  
|||||  
Db 2641 GGGAAACAGCATGGAATGGAAGAAAGTGAACCTGATGCTGATTTGCAAGATACAT 2700  
QY 2701 TCAAGGTTTCAAAAGCCGACATCTTCTGTTTCAATCCAGAAATGCAAGAGAG 2760  
|||||  
Db 2701 TCAAGGTTTCAAAAGCCGACATCTTCTGTTTCAATCCAGAAATGCAAGAGAG 2760  
QY 2761 AATGTGCAACATTTCTGCCCCTCTGAGTCTTTAAAGAAACAAAGTCCAAAGTACT 2820  
|||||  
Db 2761 AATGTGCAACATTTCTGCCCCTCTGAGTCTTTAAAGAAACAAAGTCCAAAGTACT 2820  
QY 2821 TTGAATGTGAACAAAGAGAAAGAAATCAAGAAAGATGATGATATCAAGCCGTAC 2880  
|||||  
Db 2821 TTGAATGTGAACAAAGAGAAAGAAATCAAGAAAGATGATGATATCAAGCCGTAC 2880  
QY 2881 AGACAGTAAATATACACGAGGCTTCCCTGTTGTTGTCAGAAAGATAGCCAGTTGATA 2940  
|||||  
Db 2881 AGACAGTAAATATACACGAGGCTTCCCTGTTGTTGTCAGAAAGATAGCCAGTTGATA 2940  
QY 2941 ATGCCAATGATGATCAAAAGAGGCTCTAGTGTGTTGTCATCATCTCAGTTCAAGGCA 3000  
|||||  
Db 2941 ATGCCAATGATGATCAAAAGAGGCTCTAGTGTGTTGTCATCATCTCAGTTCAAGGCA 3000  
QY 3001 ACAGAACTGAGCATATCTACCAATTAACATGAGCTTTTACAAACCCATATCGATAC 3060  
|||||  
Db 3001 ACAGAACTGAGCATATCTACCAATTAACATGAGCTTTTACAAACCCATATCGATAC 3060  
QY 3061 CACCACTTTTCCATCAAGTCAATTTGTTAAACTAATGTAAAGAAAAATCTGCTAGAG 3120  
|||||  
Db 3061 CACCACTTTTCCATCAAGTCAATTTGTTAAACTAATGTAAAGAAAAATCTGCTAGAG 3120  
QY 3121 AAAACTTTGAGGAACATCAATCTCACTGAAAGAGAAATGGAATGGAACATTTCCAA 3180  
|||||  
Db 3121 AAAACTTTGAGGAACATCAATCTCACTGAAAGAGAAATGGAATGGAACATTTCCAA 3180  
QY 3181 GTACAGTGAACCAATTAAGCCGTAATACATTTGAGAAATGTTTAAAGAGCCAGCT 3240  
|||||  
Db 3181 GTACAGTGAACCAATTAAGCCGTAATACATTTGAGAAATGTTTAAAGAGCCAGCT 3240  
QY 3241 CAAGCAATATTAATGAAGTGAAGTTCAGATCAATGAAGTGGGCTCCAGATTAATGAAA 3300  
|||||  
Db 3241 CAAGCAATATTAATGAAGTGAAGTTCAGATCAATGAAGTGGGCTCCAGATTAATGAAA 3300  
QY 3301 TAGGTTCCAGTGAATGAAGAAATTCAGAGCAAGAACTAGTGAAGAAACAGAGGCCAAATTTGA 3360  
|||||  
Db 3301 TAGGTTCCAGTGAATGAAGAAATTCAGAGCAAGAACTAGTGAAGAAACAGAGGCCAAATTTGA 3360  
QY 3361 ATGCTATGCTTACATTAAGGCTTTTCAACCTGAGGCTATTAACAAAGTCTTCCGGA 3420  
|||||  
Db 3361 ATGCTATGCTTACATTAAGGCTTTTCAACCTGAGGCTATTAACAAAGTCTTCCGGA 3420  
QY 3421 GTAATGTGAATCTCTGAAATTAAGAAAGCAAGAAATATGAAGAGTATGACAGCTGTTA 3480  
|||||



```
|||||
Db 3421 GTAATGTAGATCCTGAAATTAAGAAAGCAAGATATGAGAAAGTACTGAGCTGTA 3480
OY 3481 ATACAGATTTCTCCCATATCTGATTTAGATTAAGTGAACAGCTTATGGAGTACTG 3540
Db 3481 ATACGATTTCTCCCATATCTGATTTAGATTAAGTGAACAGCTTATGGAGTACTG 3540
OY 3541 ATGCATTCAGGTTTGTCTGTGAGACACCTGATGACCTTTAGATGATGATGATGAAATTAAGG 3600
Db 3541 ATGCATTCAGGTTTGTCTGTGAGACACCTGATGACCTTTAGATGATGATGATGAAATTAAGG 3600
OY 3601 AAGATCTAGTCTTGTCTGTAAGATGATTAAGGAAAGTCTGCTGTTTATGCAAAAGCG 3660
Db 3601 AAGATCTAGTCTTGTCTGTAAGATGATTAAGGAAAGTCTGCTGTTTATGCAAAAGCG 3660
OY 3661 TCCAGAAAGAGAGCTTGAAGAGAGTCTAGACCTTTTACCATTACATTTGGCTCAGG 3720
Db 3661 TCCAGAAAGAGAGCTTGAAGAGAGTCTAGACCTTTTACCATTACATTTGGCTCAGG 3720
OY 3721 GTTACCGAAGAGGGGCCAAGAAATAGAGTCTCAGAGAGAACTTATCTAGTGAAGT 3780
Db 3721 GTTACCGAAGAGGGGCCAAGAAATAGAGTCTCAGAGAGAACTTATCTAGTGAAGT 3780
OY 3781 AAGACCTTCCCTGCTTCCACACCTTGTATTTGTTAAAGTAAACAATATACCTTCAGT 3840
Db 3781 AAGACCTTCCCTGCTTCCACACCTTGTATTTGTTAAAGTAAACAATATACCTTCAGT 3840
OY 3841 CTACAGGCATAGACACCTGCTGACCGAGTGTCTCTTAAGAACACAGAGAAATTTAT 3900
Db 3841 CTACAGGCATAGACACCTGCTGACCGAGTGTCTCTTAAGAACACAGAGAAATTTAT 3900
OY 3901 TATCATTTGAAGAATAGTTAATGATGATGATGATGATGATGATGATGATGATGATG 3960
Db 3901 TATCATTTGAAGAATAGTTAATGATGATGATGATGATGATGATGATGATGATGATG 3960
OY 3961 AGGAACATCACCTTATGAGAGAAACAAATGTTCTGATGCTGATGCTGATGCTGATG 4020
Db 3961 AGGAACATCACCTTATGAGAGAAACAAATGTTCTGATGCTGATGCTGATGCTGATG 4020
OY 4021 GTGATTTGGAAGACTGTGACCTGCAATACAAACACCCAGAGATCTTTCTGATGCTT 4080
Db 4021 GTGATTTGGAAGACTGTGACCTGCAATACAAACACCCAGAGATCTTTCTGATGCTT 4080
OY 4081 CCAAAACAATAGAGCATAGTCTGAAAGCCAGAGAGTGTGCTGATGATGATGATGATG 4140
Db 4081 CCAAAACAATAGAGCATAGTCTGAAAGCCAGAGAGTGTGCTGATGATGATGATGATG 4140
OY 4141 TTTGAGATGATGAAGAAAGAGAGAGGCTTGAAGAAATATATCAAGAGAGCAACCA 4200
Db 4141 TTTGAGATGATGAAGAAAGAGAGAGGCTTGAAGAAATATATCAAGAGAGCAACCA 4200
OY 4201 TGGATTTCAAACTTGAAGTGAAGAGAGATCTGGGTGAGAGTGAAGAAAGAGGCTT 4260
Db 4201 TGGATTTCAAACTTGAAGTGAAGAGAGATCTGGGTGAGAGTGAAGAAAGAGGCTT 4260
OY 4261 ACTGCTCAGGCTATCTCTCAGAGTGAATTTTAACTCAGAGAGAGGATACCATG 4320
Db 4261 ACTGCTCAGGCTATCTCTCAGAGTGAATTTTAACTCAGAGAGAGGATACCATG 4320
OY 4321 AACATTAACCTGATTAAGCTCCAGAGAGAAATGGCTGAAGTGAAGTGTGTTAGAACGC 4380
Db 4321 AACATTAACCTGATTAAGCTCCAGAGAGAAATGGCTGAAGTGAAGTGTGTTAGAACGC 4380
OY 4381 ATGGAGAGCAGCTTCTTAACAGTACCTTCCATCATTAAGTACTTCTGCTGAGG 4440
Db 4381 ATGGAGAGCAGCTTCTTAACAGTACCTTCCATCATTAAGTACTTCTGCTGAGG 4440
OY 4441 ACCTGCGAAATCCAGAACAAAGACATCAGAAAAAGAGTATTAATCTCAGAGAAAGTA 4500
Db 4441 ACCTGCGAAATCCAGAACAAAGACATCAGAAAAAGAGTATTAATCTCAGAGAAAGTA 4500
OY 4501 GTGAATACCTTAAGAGCCAGAAATCCAGAAAGGCTTCTGCTGACAAAGTTGAGGTGCTG 4560
Db 4501 GTGAATACCTTAAGAGCCAGAAATCCAGAAAGGCTTCTGCTGACAAAGTTGAGGTGCTG 4560

|||||
Db 4501 GTGAATACCTTAAGAGCCAGAAATCCAGAAAGGCTTCTGCTGACAAAGTTGAGGTGCTG 4560
OY 4561 CAGATAGTCTTAACAGTAAAAATTAAGAACAGAGAGTGAAGAGTATCCCTTAAAT 4620
Db 4561 CAGATAGTCTTAACAGTAAAAATTAAGAACAGAGAGTGAAGAGTATCCCTTAAAT 4620
OY 4621 GCCCATCATTAAGATAGTGTGATGATGATGATGATGATGATGATGATGATGATG 4680
Db 4621 GCCCATCATTAAGATAGTGTGATGATGATGATGATGATGATGATGATGATGATG 4680
OY 4681 ACTACCATCTCAAGAGAGCTCATTAAGGTTTGAATGTGAGAGAGCAACAGCTGAAG 4740
Db 4681 ACTACCATCTCAAGAGAGCTCATTAAGGTTTGAATGTGAGAGAGCAACAGCTGAAG 4740
OY 4741 AGTGTGGGCGACAGCATTTGAGAGAAACATCTTACTTGGCAAGGCAAGTCTAGAGGAA 4800
Db 4741 AGTGTGGGCGACAGCATTTGAGAGAAACATCTTACTTGGCAAGGCAAGTCTAGAGGAA 4800
OY 4801 CCCCTTACCTGGAATCTGGAATCAGCCTCTCTGATGATGATGATGATGATGATGATG 4860
Db 4801 CCCCTTACCTGGAATCTGGAATCAGCCTCTCTGATGATGATGATGATGATGATGATG 4860
OY 4861 AAGACAGAGCCCGAGAGTCAAGTCTGTTGGCAACATACCATCTTCAACCTCTGATTGA 4920
Db 4861 AAGACAGAGCCCGAGAGTCAAGTCTGTTGGCAACATACCATCTTCAACCTCTGATTGA 4920
OY 4921 AAGTCCCCCAATTGGAAGTGTGGAATCTGCCAAGTCCACCTCTGCTACTACTG 4980
Db 4921 AAGTCCCCCAATTGGAAGTGTGGAATCTGCCAAGTCCACCTCTGCTACTACTG 4980
OY 4981 ATACGCTGGGTATATGCAATGGAAGAAAGTGTGAGAGAGAGAGAGAGAGAGAGAG 5040
Db 4981 ATACGCTGGGTATATGCAATGGAAGAAAGTGTGAGAGAGAGAGAGAGAGAGAGAG 5040
OY 5041 CTTCACAGAAAGGCTCAACAAAGATGTCCATGCTGCTGCTGCTGCTGCTGCTGCTG 5100
Db 5041 CTTCACAGAAAGGCTCAACAAAGATGTCCATGCTGCTGCTGCTGCTGCTGCTGCTG 5100
OY 5101 AATTTATGCTGCTGTGACAGTTTGGCAGAAACACCATCATCTTAACTTAATTA 5160
Db 5101 AATTTATGCTGCTGTGACAGTTTGGCAGAAACACCATCATCTTAACTTAATTA 5160
OY 5161 CTGAGAGACTACTATCTTGTATGAAGAAACAGATGATGATGATGATGATGATGATG 5220
Db 5161 CTGAGAGACTACTATCTTGTATGAAGAAACAGATGATGATGATGATGATGATGATG 5220
OY 5221 TGAATATTTTCTAGGAATTTGGGGAGGAAATGGGTAGTAACTATTTCTGGGTGACCC 5280
Db 5221 TGAATATTTTCTAGGAATTTGGGGAGGAAATGGGTAGTAACTATTTCTGGGTGACCC 5280
OY 5281 AGTCTATTAAAGAAAGAAATGCTGAATGAGCATGATTTTGAAGTCAAGAGAGATG 5340
Db 5281 AGTCTATTAAAGAAAGAAATGCTGAATGAGCATGATTTTGAAGTCAAGAGAGATG 5340
OY 5341 TCAATGAAGAAACCCAGAGTCCAAAGCGAGAGAGAAATCCAGAGAGAGAGATCT 5400
Db 5341 TCAATGAAGAAACCCAGAGTCCAAAGCGAGAGAGAAATCCAGAGAGAGAGATCT 5400
OY 5401 TCAGGGGCTAGAAATCTGTTGCTATGGGCCCTTACCAACATGCCCAGATCAACTGG 5460
Db 5401 TCAGGGGCTAGAAATCTGTTGCTATGGGCCCTTACCAACATGCCCAGATCAACTGG 5460
OY 5461 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5520
Db 5461 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5520
OY 5521 GCACAGTGTCCACCATTTGTTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5580
Db 5521 GCACAGTGTGTCCACCATTTGTTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5580
OY 5581 TCATGATCAATTTGGGAGATGTGTGAGGACCTGTGAGAGAGAGAGAGAGAGAGAGAG 5640
Db 5581 TCATGATCAATTTGGGAGATGTGTGAGGACCTGTGAGAGAGAGAGAGAGAGAGAGAG 5640
```

QY 5641 GTGTACACTCTACACAGTGGCCAGAGCTGACACACTTATGCCAGATCCGCCACA 5700  
|||||  
Db 5641 GTGTACACTCTACACAGTGGCCAGAGCTGACACACTTATGCCAGATCCGCCACA 5700  
QY 5701 GCCACTACTGA 5711  
|||||  
Db 5701 GCCACTACTGA 5711

RESULT 5  
AR070223 5712 bp DNA linear PAT 18-FEB-2000  
LOCUS AR070223  
DEFINITION Sequence 1 from patent US 5891857.  
ACCESSION AR070223  
VERSION AR070223.1 GI:7221111  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
1 (bases 1 to 5712)  
Holt,J.T., Jensen,R.A., King,M.-C., Page,D.L., Szabo,C.I.,  
Jettison,T.L., Robinson-Benlion,C.L. and Thompson,M.E.  
Characterized BRCA1 and BRCA2 proteins and screening and  
therapeutic methods based on characterized BRCA1 and BRCA2 proteins  
Patent: US 5891857-A 1 06-APR-1999;  
Location/Qualifiers  
1. 5712  
source 1. organism="unknown"  
BASE COUNT 1956 a 1099 c 1274 g 1383 t  
ORIGIN

Query Match 99.9%; Score 5707.8; DB 6; Length 5712;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 5709; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACTGCGCTAGACTTCTCGACCCGCCACAGCGCTGTGGGTTTCTCAGATACTGGCC 60  
|||||  
Db 1 ACTGCGCTAGACTTCTCGACCCGCCACAGCGCTGTGGGTTTCTCAGATACTGGCC 60  
QY 61 CCTGCGCTAGAGAGCGCTTACCCCTGCTGTGGTAAAGTTTCAATGGAACAGAAAGA 120  
|||||  
Db 61 CCTGCGCTAGAGAGCGCTTACCCCTGCTGTGGTAAAGTTTCAATGGAACAGAAAGA 120  
QY 121 TGGATTATCTGCTCTTGGCGTTGAAGAAGTACAAAATGTCATTAACTGATGCGAGAAA 180  
|||||  
Db 121 TGGATTATCTGCTCTTGGCGTTGAAGAAGTACAAAATGTCATTAACTGATGCGAGAAA 180  
QY 121 TGGATTATCTGCTCTTGGCGTTGAAGAAGTACAAAATGTCATTAACTGATGCGAGAAA 180  
|||||  
Db 121 TGGATTATCTGCTCTTGGCGTTGAAGAAGTACAAAATGTCATTAACTGATGCGAGAAA 180  
QY 181 TCTTAGAGTGTCCCATCTGTGAGAGTGTATCAAGAACTGTCTCCACAAAGTGTGACC 240  
|||||  
Db 181 TCTTAGAGTGTCCCATCTGTGAGAGTGTATCAAGAACTGTCTCCACAAAGTGTGACC 240  
QY 241 ACATATTTTGGCAAAATTTTGCATGCTGAACCTTCAACAGAGAAAGGCGCTTTCACAGT 300  
|||||  
Db 241 ACATATTTTGGCAAAATTTTGCATGCTGAACCTTCAACAGAGAAAGGCGCTTTCACAGT 300  
QY 301 GTCCTTATGTAAGAAATGATATACCAAAAGAGCGCTACAAAGAGTACAGAGATTAGTC 360  
|||||  
Db 301 GTCCTTATGTAAGAAATGATATACCAAAAGAGCGCTACAAAGAGTACAGAGATTAGTC 360  
QY 361 AACTGTGTAAGAGCATATGAAATATTTGCTTTTACGCTTGACACAGGTTGGAGT 420  
|||||  
Db 361 AACTGTGTAAGAGCATATGAAATATTTGCTTTTACGCTTGACACAGGTTGGAGT 420  
QY 421 ATGCAAAAGCTATATATTTTGGCAAAAAGAAATATCTCTGTAACATCTAAAGATG 480  
|||||  
Db 421 ATGCAAAAGCTATATATTTTGGCAAAAAGAAATATCTCTGTAACATCTAAAGATG 480  
QY 481 AAGTTTCTATCATCCAAAGTATGGGCTACAGAAACCGTGCCAAAGACATTCTACAGAGT 540  
|||||  
Db 481 AAGTTTCTATCATCCAAAGTATGGGCTACAGAAACCGTGCCAAAGACATTCTACAGAGT 540  
QY 541 AACCGAAAATCTCTCTTGGCAAGAAACAGTCTCAGTCCAACTCTTAACCTTGAA 600  
|||||

|||||  
Db 541 AACCGAAAATCTCTCTTGGCAAGAAACAGTCTCAGTCCAACTCTTAACCTTGAA 600  
QY 601 CTGTGACAACTCTGAGACAAAGAGCGGATACAACTCAAAAGAGCTGTCTACATTG 660  
|||||  
Db 601 CTGTGACAACTCTGAGACAAAGAGCGGATACAACTCAAAAGAGCTGTCTACATTG 660  
QY 661 AATTGGATCTGATCTTCTGAAAGATACCTTATAAGCACTTATTCAGTGTGGAG 720  
|||||  
Db 661 AATTGGATCTGATCTTCTGAAAGATACCTTATAAGCACTTATTCAGTGTGGAG 720  
QY 721 ATCAGAAATTTGTAACAATACCCCTCAAGAAACAGGAGTGAATCACTTGGATTCTG 780  
|||||  
Db 721 ATCAGAAATTTGTAACAATACCCCTCAAGAAACAGGAGTGAATCACTTGGATTCTG 780  
QY 781 CAAAAAGGCTGCTGTGAAATTTTCTGAGCGGATGTAACTGAATCACTATCAAC 840  
|||||  
Db 781 CAAAAAGGCTGCTGTGAAATTTTCTGAGCGGATGTAACTGAATCACTATCAAC 840  
QY 841 CCAATTAATTAATGATTTGAACACACAGAGCGTCAAGTGAAGGATCCAGAAAGT 900  
|||||  
Db 841 CCAATTAATTAATGATTTGAACACACAGAGCGTCAAGTGAAGGATCCAGAAAGT 900  
QY 901 ATCAGGATGCTGCTGTTTCAAACTGCAATGTGAGCGCATGTGCAAAATCACTATCA 960  
|||||  
Db 901 ATCAGGATGCTGCTGTTTCAAACTGCAATGTGAGCGCATGTGCAAAATCACTATCA 960  
QY 961 GCTCATTAACAGCATGAAACAGAGATTTTACTCATTAAGCAACAAATGATGTA 1020  
|||||  
Db 961 GCTCATTAACAGCATGAAACAGAGATTTTACTCATTAAGCAACAAATGATGTA 1020  
QY 1021 AGGCTGAATTCGTATTAATAAGCAACAGCGCTGTACAGAGGATCCAAATCACTAT 1080  
|||||  
Db 1021 AGGCTGAATTCGTATTAATAAGCAACAGCGCTGTACAGAGGATCCAAATCACTAT 1080  
QY 1081 GGGCTGGAAGTAAAGAAACATGTAATGATAGCGGATCCCAAGACAGAAAAAGGTAG 1140  
|||||  
Db 1081 GGGCTGGAAGTAAAGAAACATGTAATGATAGCGGATCCCAAGACAGAAAAAGGTAG 1140  
QY 1141 ATCTGAATGCTGATCCCTGTGTGAGAGAAAGATGAAAGCAAGCAATGCTAT 1200  
|||||  
Db 1141 ATCTGAATGCTGATCCCTGTGTGAGAGAAAGATGAAAGCAAGCAATGCTAT 1200  
QY 1201 CAGAGAAATCTAGAGATGATGAGATGCTCTGATTAACACTTAATGACAGATTGCA 1260  
|||||  
Db 1201 CAGAGAAATCTAGAGATGATGAGATGCTCTGATTAACACTTAATGACAGATTGCA 1260  
QY 1261 AAGTTAATGATGCTTTTCCAGAGATGATGATGATGATGATGATGATGATG 1320  
|||||  
Db 1261 AAGTTAATGATGCTTTTCCAGAGATGATGATGATGATGATGATGATGATG 1320  
QY 1321 GGGAGTCTGAATCAAAATGCCAAAGTATGATGATGATGATGATGATGATGATG 1380  
|||||  
Db 1321 GGGAGTCTGAATCAAAATGCCAAAGTATGATGATGATGATGATGATGATGATG 1380  
QY 1381 AATATTTCTGCTTCTGAGAGAAATAGACTTACTGCGAGTATGCTCTGAGCTTTAA 1440  
|||||  
Db 1381 AATATTTCTGCTTCTGAGAGAAATAGACTTACTGCGAGTATGCTCTGAGCTTTAA 1440  
QY 1441 TATGTAAAGTGAAGAGTTTCACTCAATTCAGTATGATGATGATGATGATGATG 1500  
|||||  
Db 1441 TATGTAAAGTGAAGAGTTTCACTCAATTCAGTATGATGATGATGATGATGATG 1500  
QY 1501 TTGGGAAAACCTATGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1560  
|||||  
Db 1501 TTGGGAAAACCTATGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1560  
QY 1561 TAATTTAGAGAGATTTTGTACTGAGCCACATATATACAGAGGCTCCCTCACAATA 1620  
|||||  
Db 1561 TAATTTAGAGAGATTTTGTACTGAGCCACATATATACAGAGGCTCCCTCACAATA 1620  
QY 1621 AATTAAGCGTAAAGAGACCTTACATCAGGCTTATCTGAGGATTTTATCAAGAA 1680  
|||||

Dh 1621 AATTAAAGCGTAAAGAGACCTACATCGGCTTTCATCTGAGATTTCATCAAGAAAG 1680  
Qy 1681 CAGATTGGCAGTTCAAAGAGACTCCTGGAATGATTAATCAGGAGACTAACCAAGCGAGC 1740  
Dh 1681 CAGATTGGCAGTTCAAAGAGACTCCTGGAATGATTAATCAGGAGACTAACCAAGCGAGC 1740  
Qy 1741 AGAATGGTCAAGTGAATATTTACTATACTGGTGCATGAGAAATAAACAAAGGTGAT 1800  
Dh 1741 AGAATGGTCAAGTGAATATTTACTATACTGGTGCATGAGAAATAAACAAAGGTGAT 1800  
Qy 1801 CTATTCAAGATGAGAAAAATCTTAACCAATAGATCATCTGAAAAAGATCTGTTTCA 1860  
Dh 1801 CTATTCAAGATGAGAAAAATCTTAACCAATAGATCATCTGAAAAAGATCTGTTTCA 1860  
Qy 1861 AAACGAAGCTGAACCTTAAGCAGCAGTATAGCAATATGGAATGGAATGGAATATATCC 1920  
Dh 1861 AAACGAAGCTGAACCTTAAGCAGCAGTATAGCAATATGGAATGGAATGGAATATATCC 1920  
Qy 1921 ACAATTCAAAAGCAGCTTAAGAAATAGGCTGAGGAGAGAGCTTCTACAGGCAATATTC 1980  
Dh 1921 ACAATTCAAAAGCAGCTTAAGAAATAGGCTGAGGAGAGAGCTTCTACAGGCAATATTC 1980  
Qy 1981 ATGGCTTGAACCTGATGATGCTAGTAAATCTAAGCCACCTAATTTGATGATTCGAAA 2040  
Dh 1981 ATGGCTTGAACCTGATGATGCTAGTAAATCTAAGCCACCTAATTTGATGATTCGAAA 2040  
Qy 2041 TTGATAGTGTTCAGAGTGAAGATTAAGAAAAAATGATACACCAATGGCAGTCA 2100  
Dh 2041 TTGATAGTGTTCAGAGTGAAGATTAAGAAAAAATGATACACCAATGGCAGTCA 2100  
Qy 2101 GGCACAGCAGAAACCTCAACATCATGGAAGTAAAGAACTGCACTGAGCAGCAAGAAGA 2160  
Dh 2101 GGCACAGCAGAAACCTCAACATCATGGAAGTAAAGAACTGCACTGAGCAGCAAGAAGA 2160  
Qy 2161 GTAACAAAGCCAAATGAGACAGACAGATTAAGAGATGACATGATTTTCCAGAGCTGA 2220  
Dh 2161 GTAACAAAGCCAAATGAGACAGACAGATTAAGAGATGACATGATTTTCCAGAGCTGA 2220  
Qy 2221 AGTAAACAAATGACAGCTGTTCTTTACTAGTGTCAATATGACAGTAACTTAAAGAT 2280  
Dh 2221 AGTAAACAAATGACAGCTGTTCTTTACTAGTGTCAATATGACAGTAACTTAAAGAT 2280  
Qy 2281 TTGTCAATCTTACCTTCCAGAGAGAGAAAAAGAGAACTAGAAAACATTTAAAGTGT 2340  
Dh 2281 TTGTCAATCTTACCTTCCAGAGAGAGAAAAAGAGAACTAGAAAACATTTAAAGTGT 2340  
Qy 2341 CTAAATATGCTGAGAGACCCCAAGATCTCATGTTAGTGAGAAAGGTTTTGCAAACTG 2400  
Dh 2341 CTAAATATGCTGAGAGACCCCAAGATCTCATGTTAGTGAGAAAGGTTTTGCAAACTG 2400  
Qy 2401 AAAGATCTGAGAGATGAGATTTTCATTTGTTACCTGGTACGATTTATGAGCACTGAG 2460  
Dh 2401 AAAGATCTGAGAGATGAGATTTTCATTTGTTACCTGGTACGATTTATGAGCACTGAG 2460  
Qy 2461 AAAGATCTGAGATGAGATTTTCATTTGTTACCTGGTACGATTTATGAGCACTGAG 2520  
Dh 2461 AAAGATCTGAGATGAGATTTTCATTTGTTACCTGGTACGATTTATGAGCACTGAG 2520  
Qy 2521 GTGTGAGTCACTGTGAGACATTTTAAACCAGGAGCTAATTTGATGTTTCCAAAG 2580  
Dh 2521 GTGTGAGTCACTGTGAGACATTTTAAACCAGGAGCTAATTTGATGTTTCCAAAG 2580  
Qy 2581 ATATATGAATGAGACAGAGGCTTTAGTATTCATTCATGAGGAGCTAATTTGATGTTTCC 2640  
Dh 2581 ATATATGAATGAGACAGAGGCTTTAGTATTCATTCATGAGGAGCTAATTTGATGTTTCC 2640  
Qy 2641 GGGAAACAGAGCTGAATGAGAGAGAGTGAATTTGATGTTTCCAAAG 2700  
Dh 2641 GGGAAACAGAGCTGAATGAGAGAGAGTGAATTTGATGTTTCCAAAG 2700  
Qy 2701 TCAAGGTTTCAAGGCGCAGTCAATTTGCTGTTTCAATTCAGAAATGACAGAGAG 2760  
Dh 2701 TCAAGGTTTCAAGGCGCAGTCAATTTGCTGTTTCAATTCAGAAATGACAGAGAG 2760

Qy 2761 AATGTGCAACATTCCTCTGCCCCACCTGCTGCTTTAAAGAAACAAAGTCCAAAGTCACTT 2820  
Dh 2761 AATGTGCAACATTCCTCTGCCCCACCTGCTGCTTTAAAGAAACAAAGTCCAAAGTCACTT 2820  
Qy 2821 TTGAATGTGAACAAAGAGAGAAAAATCAAGAGAAATGAGCTTAATATCAAGCCTGTAC 2880  
Dh 2821 TTGAATGTGAACAAAGAGAGAAAAATCAAGAGAAATGAGCTTAATATCAAGCCTGTAC 2880  
Qy 2881 AGACAGTTAATATCACTGAGGCTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2940  
Dh 2881 AGACAGTTAATATCACTGAGGCTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2940  
Qy 2941 ATGCCAAATGTAGATCAAGAGAGGCTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3000  
Dh 2941 ATGCCAAATGTAGATCAAGAGAGGCTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3000  
Qy 3001 ACAGAACTGACATTAATCTCAAAATTAACATGAGACTTTTCAAAACCCATATGCTATAC 3060  
Dh 3001 ACAGAACTGACATTAATCTCAAAATTAACATGAGACTTTTCAAAACCCATATGCTATAC 3060  
Qy 3061 CACCACTTTTCCATCAAGTCAATTTTGTAAATCAATATGTAAGAAAAATCTGCTAGAG 3120  
Dh 3061 CACCACTTTTCCATCAAGTCAATTTTGTAAATCAATATGTAAGAAAAATCTGCTAGAG 3120  
Qy 3121 AAAACTTGGAGACATTCATGATGACCTGAAAGAGAAATGGAATGGAACATTCGAA 3180  
Dh 3121 AAAACTTGGAGACATTCATGATGACCTGAAAGAGAAATGGAATGGAACATTCGAA 3180  
Qy 3181 GTACAGTGAAGCAATTAAGCCTTAATATAGAGAAATGTTTAAAGAGCAGCT 3240  
Dh 3181 GTACAGTGAAGCAATTAAGCCTTAATATAGAGAAATGTTTAAAGAGCAGCT 3240  
Qy 3241 CAAGCAATATTAATGAGTACGTTCCAGTACTAATGAAGTGGCTCCAGTATTAATGAAA 3300  
Dh 3241 CAAGCAATATTAATGAGTACGTTCCAGTACTAATGAAGTGGCTCCAGTATTAATGAAA 3300  
Qy 3301 TAGGTTCCAGTGAAGAAACATTCAGCAGAGAACTAGTGAAGAAACAGAGGCGCAAAATGGA 3360  
Dh 3301 TAGGTTCCAGTGAAGAAACATTCAGCAGAGAACTAGTGAAGAAACAGAGGCGCAAAATGGA 3360  
Qy 3361 ATGCTATGCTTAGATTAAGGCTTTTGAACCTGAGGCTTATTAACAAAGCTTCTCGGAA 3420  
Dh 3361 ATGCTATGCTTAGATTAAGGCTTTTGAACCTGAGGCTTATTAACAAAGCTTCTCGGAA 3420  
Qy 3421 GTAATTTGAGCATCTCTAATTAAGAAAGCAAGATATGAAGAGTACTAGACTGTTA 3480  
Dh 3421 GTAATTTGAGCATCTCTAATTAAGAAAGCAAGATATGAAGAGTACTAGACTGTTA 3480  
Qy 3481 ATACAGATTTCTGTCATATCTGATTTGATTAAGATTAAGACAGCGCTATGGAGTATGC 3540  
Dh 3481 ATACAGATTTCTGTCATATCTGATTTGATTAAGATTAAGACAGCGCTATGGAGTATGC 3540  
Qy 3541 ATGCAATCTGAGGTTTGTCTGAGACACCTGATGACCTGTTAGATGATGTAATTAAGG 3600  
Dh 3541 ATGCAATCTGAGGTTTGTCTGAGACACCTGATGACCTGTTAGATGATGTAATTAAGG 3600  
Qy 3601 AAGATTAAGTGTGCTGAGAAATGACATTAAGAAAGTCTGCTGTTTAAAGCAAGCG 3660  
Dh 3601 AAGATTAAGTGTGCTGAGAAATGACATTAAGAAAGTCTGCTGTTTAAAGCAAGCG 3660  
Qy 3661 TCCAGAAAGAGAGACTTAAGCAGAGGCTTACCTGCTTACCCATTAACATTTGCTCAGG 3720  
Dh 3661 TCCAGAAAGAGAGACTTAAGCAGAGGCTTACCTGCTTACCCATTAACATTTGCTCAGG 3720  
Qy 3721 GTTACGGAAGAGGAGGAGAAATTAAGATTCCTGAGAGAGAACTTATCTAGAGATG 3780  
Dh 3721 GTTACGGAAGAGGAGGAGAAATTAAGATTCCTGAGAGAGAACTTATCTAGAGATG 3780  
Qy 3781 AAGAGCTTCCCTGCTTCCACACCTGTTATTTGTTAAAGTAAACATTAATACCTTCACT 3840  
Dh 3781 AAGAGCTTCCCTGCTTCCACACCTGTTATTTGTTAAAGTAAACATTAATACCTTCACT 3840

3841 CTAAGCATAGACACCGCTTCTACCGAGTGTCTCTAAGACACAGAGAGATTTAT 3900  
3841 CTACTAGCATAGACACCGCTTCTACCGAGTGTCTCTAAGACACAGAGAGATTTAT 3900  
3901 TATCATGAGAAATAGCTTAATATGATGACATACAGATATATATGGCAAGGATCTC 3960  
3901 TATCATGAGAAATAGCTTAATATGATGACATACAGATATATATGGCAAGGATCTC 3960  
3961 AGGAACATCACTTAGTAGAGAAACAAATGTTGCTAGCTGTTTCTTCTACAGTCA 4020  
3961 AGGAACATCACTTAGTAGAGAAACAAATGTTGCTAGCTGTTTCTTCTACAGTCA 4020  
4021 GTGATTTGGAAGACTTACTGCAAAATACAAACACCCAGATCTTCTGATTTGTT 4080  
4021 GTGATTTGGAAGACTTACTGCAAAATACAAACACCCAGATCTTCTGATTTGTT 4080  
4081 CCAACCAATAGGAGCATCAGTCTGAAGCCAGGAGTGTCTGATGATCAGAGATTTG 4140  
4081 CCAACCAATAGGAGCATCAGTCTGAAGCCAGGAGTGTGTGATGATCAGAGATTTG 4140  
4141 TTTGAGTGTAGAGAAAGAGAGACGGGCTTGAAGAAATATATCAAGAGAGCA 4200  
4141 TTTGAGTGTAGAGAAAGAGAGACGGGCTTGAAGAAATATATCAAGAGAGCA 4200  
4201 TGGATTTCAACTTATAGTGAAGACAGATCTGGGTGTGAGATGAACAAGCTCTCTG 4260  
4201 TGGATTTCAACTTATAGTGAAGACAGATCTGGGTGTGAGATGAACAAGCTCTCTG 4260  
4261 ACAGCTCAGGAGCTATCCTCTCAGAGTACATTTTACACACTCAGCAGAGGATAC 4320  
4261 ACAGCTCAGGAGCTATCCTCTCAGAGTACATTTTACACACTCAGCAGAGGATAC 4320  
4321 AACATTAAGCTGATTAAGCTCCAGCAGAAATGCTGAATAGAACTGTGTTAAGAC 4380  
4321 AACATTAAGCTGATTAAGCTCCAGCAGAAATGCTGAATAGAACTGTGTTAAGAC 4380  
4381 ATGGAGCCAGCCCTTCTAAGAGCTACCTTCCATCAATAGAGTCTTCTGCTGAGG 4440  
4381 ATGGAGCCAGCCCTTCTAAGAGCTACCTTCCATCAATAGAGTCTTCTGCTGAGG 4440  
4441 ACCTGGCAATTCAGAAACAAAGCAGATCAGAAAGAGATTAATTAATTTACAGAA 4500  
4441 ACCTGGCAATTCAGAAACAAAGCAGATCAGAAAGAGATTAATTAATTTACAGAA 4500  
4501 GTGAATTAAGCTTATAGCCAGAAATCCAGAGGCTTCTGCTGACAAAGTTGAGT 4560  
4501 GTGAATTAAGCTTATAGCCAGAAATCCAGAGGCTTCTGCTGACAAAGTTGAGT 4560  
4561 CAGATAGTCTTACAGAAATTAAGAAACAGAGAGTGAAGAGTCAATCCCTCTTAAT 4620  
4561 CAGATAGTCTTACAGAAATTAAGAAACAGAGAGTGAAGAGTCAATCCCTCTTAAT 4620  
4621 GCCCATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4680  
4621 GCCCATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4680  
4681 ACTACCATCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4740  
4681 ACTACCATCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4740  
4741 AGTCTGGGCGACAGATTTGAGGAAACATCTTACTTCAAGGAGATCTAAGAGGAA 4800  
4741 AGTCTGGGCGACAGATTTGAGGAAACATCTTACTTCAAGGAGATCTAAGAGGAA 4800  
4801 CCCCTTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4860  
4801 CCCCTTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4860  
4861 AAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4920  
4861 AAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4920  
4921 AAGTTCCCAATTTGAAGATTTGAGAGATTTGCCAGAGTCCAGCTGCTCTCATCTG 4980  
4921 AAGTTCCCAATTTGAAGATTTGAGAGATTTGCCAGAGTCCAGCTGCTCTCATCTG 4980

4921 AAGTTCCCAATTTGAAGATTTGAGAGATTTGCCAGAGTCCAGCTGCTCTCATCTG 4980  
4981 ATACTGCTGGGTATATGCAATGAGAGAAAGTGTGACAGAGAGAGAGAGAGAG 5040  
4981 ATACTGCTGGGTATATGCAATGAGAGAAAGTGTGACAGAGAGAGAGAGAGAG 5040  
5041 CTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5100  
5041 CTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5100  
5101 AATTATGCTGTGTACAGAGTTTCCGAGAAACACCATCTTAACTTAACTTAA 5160  
5101 AATTATGCTGTGTGTACAGAGTTTCCGAGAAACACCATCTTAACTTAACTTAA 5160  
5161 CTGAGAGAGACTACTCATCTGTTTATGAAACAGATGCTGTTGTTGTTGAGAG 5220  
5161 CTGAGAGAGACTACTCATCTGTTTATGAAACAGATGCTGTTGTTGTTGAGAG 5220  
5221 TGAATATATTTCTAGAAATTTGCGGAGAGAAATGCGTAGTATTTCTGGGTACC 5280  
5221 TGAATATATTTCTAGAAATTTGCGGAGAGAAATGCGTAGTATTTCTGGGTACC 5280  
5281 AGCTATTTAAAGAAAGAAATGCTGATGATGATGATGATGATGATGATGATG 5340  
5281 AGCTATTTAAAGAAAGAAATGCTGATGATGATGATGATGATGATGATGATG 5340  
5341 TCAATGAGAGAAACACCAAGGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAG 5400  
5341 TCAATGAGAGAAACACCAAGGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAG 5400  
5401 TCAGGAGGCTAGAAATGTTGCTATGAGGCTTCCACCAACATGCCACAGATCA 5460  
5401 TCAGGAGGCTAGAAATGTTGCTATGAGGCTTCCACCAACATGCCACAGATCA 5460  
5461 AATGAGTGTACAGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5520  
5461 AATGAGTGTACAGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5520  
5521 GCACAGTGTCCACCAATGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5580  
5521 GCACAGTGTCCACCAATGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5580  
5581 TCATGCAATTTGGGAGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5640  
5581 TCATGCAATTTGGGAGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5640  
5641 GTGTAGCAGCTTACCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5700  
5641 GTGTAGCAGCTTACCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5700  
5701 GCCCACTACTGA 5711  
5701 GCCCACTACTGA 5711  
5711 GCCCACTACTGA 5711  
5711 GCCCACTACTGA 5711

RESULT 6  
ARI18507 5712 bp DNA 1linear PAT 16-MAY-2001  
LOCUS ARI18507 Sequence 1 from patent US 6149903.  
DEFINITION ARI18507  
ACCESSION ARI18507  
VERSION ARI18507.1 GI:14100417  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE  
1 (bases 1 to 5712)  
Holt,J.T., Jensen,R.A., King,M.-C., Page,D.L., Szabo,C.I.,  
Jelton,T.L., Robinson-Benion,C.L. and Thompson,M.E.  
Characterized BRCA1 and BRCA2 proteins and screening and  
therapeutic methods based on characterized BRCA1 and BRCA2 proteins  
Patent: US 6149903-A 1 21-NOV-2000;  
JOURNAL Location/Qualifiers  
FEATURES

source 1. 5712  
/organism="unknown"  
BASE COUNT 1956 a 1099 c 1274 g 1383 t  
ORIGIN

Query Match 99.9%; Score 5707.8; DB 6; Length 5712;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 5709; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Oy 1 AGCTGCGTGAAGATCTCTGACCCCGACCGAGGCTGTGGGTTCTCAGATACTGGGCC 60
Db 1 AGCTGCGTGAAGATCTCTGACCCCGACCGAGGCTGTGGGTTCTCAGATACTGGGCC 60
Oy 61 CCTGCGCTCAGAGAGCCCTTCAACCCTGCTGGGTAAGTTCAATTGGAGAGAGAAA 120
Db 61 CCTGCGCTCAGAGAGCCCTTCAACCCTGCTGGGTAAGTTCAATTGGAGAGAGAAA 120
Oy 121 TGAATTTATCTGCTCTTCCGCTTGAAGAGTACAAATGTCATTATGCTATGAGAAA 180
Db 121 TGAATTTATCTGCTCTTCCGCTTGAAGAGTACAAATGTCATTATGCTATGAGAAA 180
Oy 181 TCTTAGAGTGTCCATGCTGCTGAGTTGATCAAGGAACCTGCTCCACAAAGTGTACC 240
Db 181 TCTTAGAGTGTCCATGCTGCTGAGTTGATCAAGGAACCTGCTCCACAAAGTGTACC 240
Oy 241 ACATATTTTGGCAATTTTGCATGCTGAACCTTCTCAACCAAGAGAGAGGCCCTTCA 300
Db 241 ACATATTTTGGCAATTTTGCATGCTGAACCTTCTCAACCAAGAGAGAGGCCCTTCA 300
Oy 301 GTCTTTATGTAGATGATTAACCAAAAGAGCCCTTCAAGAAAGTACAGATTTAGTC 360
Db 301 GTCTTTATGTAGATGATTAACCAAAAGAGCCCTTCAAGAAAGTACAGATTTAGTC 360
Oy 361 AACTTTGTAAGAGCTATTGAAATCATTTGTGCTTTTACGCTTACACAGGTTGAGT 420
Db 361 AACTTTGTAAGAGCTATTGAAATCATTTGTGCTTTTACGCTTACACAGGTTGAGT 420
Oy 421 ATGCAACAGCTATATTTTGCAGAAAAAGAAATTAACCTCTCAACATCTAAAGATG 480
Db 421 ATGCAACAGCTATATTTTGCAGAAAAAGAAATTAACCTCTCAACATCTAAAGATG 480
Oy 481 AAGTTTCTATCATCCAAAGTATGGGCTACAGAAACCGTCCAAAAGACTTTACAGATG 540
Db 481 AAGTTTCTATCATCCAAAGTATGGGCTACAGAAACCGTCCAAAAGACTTTACAGATG 540
Oy 541 AACCAGAAATCTCTCTGAGAGAAACAGTCTCAGTGTCCACCTCTTAACCTTGGAA 600
Db 541 AACCAGAAATCTCTCTGAGAGAAACAGTCTCAGTGTCCACCTCTTAACCTTGGAA 600
Oy 601 CTGTGAGAACTCTGAGAGCAAAAGCAGGATACAACTCAAAAAGAGCTGTACATTTG 660
Db 601 CTGTGAGAACTCTGAGAGCAAAAGCAGGATACAACTCAAAAAGAGCTGTACATTTG 660
Oy 661 AATTGGATCTGATTTCTTGAAGATACCGTTAATAAGCAACTTATGACAGTGTGGAG 720
Db 661 AATTGGATCTGATTTCTTGAAGATACCGTTAATAAGCAACTTATGACAGTGTGGAG 720
Oy 721 ATCAAGAATTGTTCAATATCACCCTTAAGAGAACAGGATGAATATGATTTGATTCG 780
Db 721 ATCAAGAATTGTTCAATATCACCCTTAAGAGAACAGGATGAATATGATTTGATTCG 780
Oy 781 CAAAAAAGGCTGTTGAAATTTCTGAGAGGATGTAAACAATATGAAACATCATCAAC 840
Db 781 CAAAAAAGGCTGTTGAAATTTCTGAGAGGATGTAAACAATATGAAACATCATCAAC 840
Oy 841 CCAATATTAATGATTTGAACACCACTGAGAGGCTGAGAGGATCCAGAAAAGT 900
Db 841 CCAATATTAATGATTTGAACACCACTGAGAGGCTGAGAGGATCCAGAAAAGT 900
Oy 901 ATCAGGAGTGTCTTCAAACTTGCATGTGAGAGCATGGGACAAATATCATGSCA 960
Db 901 ATCAGGAGTGTCTTCAAACTTGCATGTGAGAGCATGGGACAAATATCATGSCA 960

```

```

Oy 961 GCTCATTTACAGCATGAGAACAGAGATTATTTACTTCACTTAAGAGAGATGATGAGAAA 1020
Db 961 GCTCATTTACAGCATGAGAACAGAGATTATTTACTTCACTTAAGAGAGATGATGAGAAA 1020
Oy 1021 AGCGTAATTTCTGTAATTAAGCAACAGCCTGGCTTAGCAAGAGACCACATACAGAT 1080
Db 1021 AGCGTAATTTCTGTAATTAAGCAACAGCCTGGCTTAGCAAGAGACCACATACAGAT 1080
Oy 1081 GGGCTGAGAGTAAAGAAATGTAATGATAGGCGAGCTCCAGACACAGAAAAAGGTAG 1140
Db 1081 GGGCTGAGAGTAAAGAAATGTAATGATAGGCGAGCTCCAGACACAGAAAAAGGTAG 1140
Oy 1141 ATCTGAATGCTGATCCCTGTGTGAGAGAAAAAGATGAATTAAGCAGAACTGCATGCT 1200
Db 1141 ATCTGAATGCTGATCCCTGTGTGAGAGAAAAAGATGAATTAAGCAGAACTGCATGCT 1200
Oy 1201 CAGAGATCTAGAGATTAAGATGTTCTTGGATTAACATTAATAGCAGCATTCAGA 1260
Db 1201 CAGAGATCTAGAGATTAAGATGTTCTTGGATTAACATTAATAGCAGCATTCAGA 1260
Oy 1261 AAGTTAATGAGTGGTTTCCAGAGTGAATGAACTGTAGGTTGATGATGACATGATG 1320
Db 1261 AAGTTAATGAGTGGTTTCCAGAGTGAATGAACTGTAGGTTGATGATGACATGATG 1320
Oy 1321 GGGAGTCTGAATCAAAATGCCAAAGTAGCTGATTTGAGAGAGTTCTAAATAGGTAGATG 1380
Db 1321 GGGAGTCTGAATCAAAATGCCAAAGTAGCTGATTTGAGAGAGTTCTAAATAGGTAGATG 1380
Oy 1381 AATATTTCTGCTTCTTCAAGAGAAATAGACTTATGCGCAGTATCTCATAGGCTTTAA 1440
Db 1381 AATATTTCTGCTTCTTCAAGAGAAATAGACTTATGCGCAGTATCTCATAGGCTTTAA 1440
Oy 1441 TATGTAAAGTGAAGAGTTCACATCAAAATAGTAAAGATTAATGAAACAAATAT 1500
Db 1441 TATGTAAAGTGAAGAGTTCACATCAAAATAGTAAAGATTAATGAAACAAATAT 1500
Oy 1501 TTGGAAAAACCTATTCGAGAGAGCAAGCCTCCCAACTTAAGCATGTAACTGAAATC 1560
Db 1501 TTGGAAAAACCTATTCGAGAGAGCAAGCCTCCCAACTTAAGCATGTAACTGAAATC 1560
Oy 1561 TAATTTATAGAGCAATTTGTTACTAGACACAGATTAATCAAGAGCGTCCCTCAATA 1620
Db 1561 TAATTTATAGAGCAATTTGTTACTAGACACAGATTAATCAAGAGCGTCCCTCAATA 1620
Oy 1621 AATTAAAGCGTAAAGAGAGCTACATCAGGCTTCATCCGAGATTTTATCAAGAAAG 1680
Db 1621 AATTAAAGCGTAAAGAGAGCTACATCAGGCTTCATCCGAGATTTTATCAAGAAAG 1680
Oy 1681 CAGATTTGGCAGTTCAAAAGACTCCTGAATATGATAAATCAGAGAACTAACCAAGCAG 1740
Db 1681 CAGATTTGGCAGTTCAAAAGACTCCTGAATATGATAAATCAGAGAACTAACCAAGCAG 1740
Oy 1741 AGAATGCTCAAGTATGATTAATTAATAGTGTCTATGAGAAATTAAGAAAGGTGAT 1800
Db 1741 AGAATGCTCAAGTATGATTAATTAATAGTGTCTATGAGAAATTAAGAAAGGTGAT 1800
Oy 1801 CTATTCAGAAATGAGAAAAATCCTTAACCAATAGAAATACCTGGAAGAAATCTCTTCA 1860
Db 1801 CTATTCAGAAATGAGAAAAATCCTTAACCAATAGAAATACCTGGAAGAAATCTCTTCA 1860
Oy 1861 AAGAGAAAGCTGAACCTATAGAGCAGATTAAGCAATATGGAATCTGGAATTAATATCC 1920
Db 1861 AAGAGAAAGCTGAACCTATAGAGCAGATTAAGCAATATGGAATCTGGAATTAATATCC 1920
Oy 1921 ACAATTTCAAAAGCACCTTAAGAAATAGGCTGAGAGAGAGTCTTCTACACAGGATATTC 1980
Db 1921 ACAATTTCAAAAGCACCTTAAGAAATAGGCTGAGAGAGAGTCTTCTACACAGGATATTC 1980
Oy 1981 ATGCGCTTGAACCTGATGATGATGATTAACCAACCTAATTTAGTGAATTTGAAA 2040
Db 1981 ATGCGCTTGAACCTGATGATGATGATTAACCAACCTAATTTAGTGAATTTGAAA 2040
Oy 2041 TTGATAGTGTCTTAGAGAGTGAAGATTAAGAAAAAAGTACCAATATGCTCACTCA 2100

```

Db 2041 TTGATAGTTGTTCTACAGTAGAGAGATTAAGAAAAAGTACAACTAATGCCACTGCA 2100  
 QY 2101 GGCACAGCGAAACCTACACACTCATGGAAGGTAAAGAACTGCACAGTGGAGCCAGAGAGA 2160  
 Db 2101 GGCACAGCGAAACCTACACACTCATGGAAGGTAAAGAACTGCACAGTGGAGCCAGAGAGA 2160  
 QY 2161 GTAAACAGCCAAATGAACAGACAGTAAAGACATGACAGTGAATCTTCCAGAGCTGA 2220  
 Db 2161 GTAAACAGCCAAATGAACAGACAGTAAAGACATGACAGTGAATCTTCCAGAGCTGA 2220  
 QY 2221 AGTTAAACAATGCACTGGTCTTTTACTAAGTGTCAATACCAGTGAACCTTAAGAAAT 2280  
 Db 2221 AGTTAAACAATGCACTGGTCTTTTACTAAGTGTCAATACCAGTGAACCTTAAGAAAT 2280  
 QY 2281 TTGTCAATCCTTACCTTCCAGAGAGAAAAAGAGAACTGAACAACTTAAGTGT 2340  
 Db 2281 TTGTCAATCCTTACCTTCCAGAGAGAAAAAGAGAACTGAACAACTTAAGTGT 2340  
 QY 2341 CTAAATATGCTGAAGACCCCAAGATCTCATGTTAACTGAGAAAGGGTTTGCACACTG 2400  
 Db 2341 CTAAATATGCTGAAGACCCCAAGATCTCATGTTAACTGAGAAAGGGTTTGCACACTG 2400  
 QY 2401 AAAGATCTGTAGAGAGTAGCATATTTTCATTTGCTGACTGATTTATGGCAGTCAAG 2460  
 Db 2401 AAAGATCTGTAGAGAGTAGCATATTTTCATTTGCTGACTGATTTATGGCAGTCAAG 2460  
 QY 2461 AAAGATCTGTAGAGAGTAGCATATTTTCATTTGCTGACTGATTTATGGCAGTCAAG 2520  
 Db 2461 AAAGATCTGTAGAGAGTAGCATATTTTCATTTGCTGACTGATTTATGGCAGTCAAG 2520  
 QY 2521 GTGTGAGTCACTGTGAGCATTTTGAAGAACCCCAAGGACTAATTCATGTTGTTCACAG 2580  
 Db 2521 GTGTGAGTCACTGTGAGCATTTTGAAGAACCCCAAGGACTAATTCATGTTGTTCACAG 2580  
 QY 2581 ATAAATGAAGATGACACAGAAAGGCTTTAATATCCATTGGACATGAAGTTAACACAGTC 2640  
 Db 2581 ATAAATGAAGATGACACAGAAAGGCTTTAATATCCATTGGACATGAAGTTAACACAGTC 2640  
 QY 2641 GGGAAACAAGCATAGAAATGGAAGAAAGTGAATGCTGCTCAATTTTCAGAAATTCAT 2700  
 Db 2641 GGGAAACAAGCATAGAAATGGAAGAAAGTGAATGCTGCTCAATTTTCAGAAATTCAT 2700  
 QY 2701 TCAAGGTTTCAAGGCCGACGTATTTGCTGTTTCMAATCCAGAAATGCAGAAAGAG 2760  
 Db 2701 TCAAGGTTTCAAGGCCGACGTATTTGCTGTTTCMAATCCAGAAATGCAGAAAGAG 2760  
 QY 2761 AATGTGCAACATTTCTGCCCACCTGTGGTCTTTAAAGAAACAAGTCCAAAAGTCACTT 2820  
 Db 2761 AATGTGCAACATTTCTGCCCACCTGTGGTCTTTAAAGAAACAAGTCCAAAAGTCACTT 2820  
 QY 2821 TTGAATGTGAACAAAAGAAAGAAATCAAGAAAGATGAGTCAATATCAGGCTGTAC 2880  
 Db 2821 TTGAATGTGAACAAAAGAAAGAAATCAAGAAAGATGAGTCAATATCAGGCTGTAC 2880  
 QY 2881 AGACAGTTAATATCACTGCAAGGCTTCTGTGTGTGTGTCAGAAAGATTAAGCCAGTTGATA 2940  
 Db 2881 AGACAGTTAATATCACTGCAAGGCTTCTGTGTGTGTGTCAGAAAGATTAAGCCAGTTGATA 2940  
 QY 2941 ATGCCAATGTAGTATCAAGAGAGGCTTAGGTTTGTCTATCATCTCACTTCAGAGGCA 3000  
 Db 2941 ATGCCAATGTAGTATCAAGAGAGGCTTAGGTTTGTCTATCATCTCACTTCAGAGGCA 3000  
 QY 3001 AGGAAACTGGACATTTACTCCAAATTAACATGAGACTTTTACAAAACCCATATCGTATAC 3060  
 Db 3001 AGGAAACTGGACATTTACTCCAAATTAACATGAGACTTTTACAAAACCCATATCGTATAC 3060  
 QY 3061 CACCACTTTTCCCATCAAGTCAATTTGTTAAACTAAATGTAAAGAAAATCTCTAGAGG 3120  
 Db 3061 CACCACTTTTCCCATCAAGTCAATTTGTTAAACTAAATGTAAAGAAAATCTCTAGAGG 3120  
 QY 3121 AAAACTTTGAGAACATTTCAATGTCACTGAAAGAGAAATGGGAAATGAGACATTTCCAA 3180  
 Db 3121 AAAACTTTGAGAACATTTCAATGTCACTGAAAGAGAAATGGGAAATGAGACATTTCCAA 3180

Db 3121 AAAACTTTGAGAACATTTCAATGTCACTGAAAGAGAAATGGGAAATGAGACATTTCCAA 3180  
 QY 3181 GTACAGTGAGACACATTTAGCCGTAATATACATTTAGAGAAAATGTTTTAAAGAACGCCAGCT 3240  
 Db 3181 GTACAGTGAGACACATTTAGCCGTAATATACATTTAGAGAAAATGTTTTAAAGAACGCCAGCT 3240  
 QY 3241 CAAGCAATATTAATGAAGTAGTGTCCAGTACTAATGAAGTGGGCTCCAGTATTTAATGA 3300  
 Db 3241 CAAGCAATATTAATGAAGTAGTGTCCAGTACTAATGAAGTGGGCTCCAGTATTTAATGA 3300  
 QY 3301 TAGGTTCCAGTATGAAGAACATTTACAGCAGAACTAGTATGAAGAACAGAGGCCAAATTTGA 3360  
 Db 3301 TAGGTTCCAGTATGAAGAACATTTACAGCAGAACTAGTATGAAGAACAGAGGCCAAATTTGA 3360  
 QY 3361 ATGCTATGCTTAGATTTAGGGGTTTGTGCAACCTAGGCTCTTAACAAAGTCTCTCGAA 3420  
 Db 3361 ATGCTATGCTTAGATTTAGGGGTTTGTGCAACCTAGGCTCTTAACAAAGTCTCTCGAA 3420  
 QY 3421 GTAAATTTAGCAATCTGAAATTAAGAAAGCAGAAATATGAAGATAGTTTCAGACTGTGA 3480  
 Db 3421 GTAAATTTAGCAATCTGAAATTAAGAAAGCAGAAATATGAAGATAGTTTCAGACTGTGA 3480  
 QY 3481 ATACAGATTTCTCTCCATATCTGATTTACATATCTAGATATCTAGACAGCTATGAGTGC 3540  
 Db 3481 ATACAGATTTCTCTCCATATCTGATTTACATATCTAGATATCTAGACAGCTATGAGTGC 3540  
 QY 3541 ATGCATCTCAGGTTTGTCTGAGACACCTGATACCTGTTAGATGATGATGAATTAAGG 3600  
 Db 3541 ATGCATCTCAGGTTTGTCTGAGACACCTGATACCTGTTAGATGATGATGAATTAAGG 3600  
 QY 3601 AAGATACTAGTTTGTGCGAAGAAATGACATTAAGGAAAGTCTGCTGTTTACCAAAAGCG 3660  
 Db 3601 AAGATACTAGTTTGTGCGAAGAAATGACATTAAGGAAAGTCTGCTGTTTACCAAAAGCG 3660  
 QY 3661 TCCAGAAAGAGAGCTTAGCAGAGAGTCTAGCCCTTTACACCATACACATTTGGCTCAGG 3720  
 Db 3661 TCCAGAAAGAGAGCTTAGCAGAGAGTCTAGCCCTTTACACCATACACATTTGGCTCAGG 3720  
 QY 3721 GTTACCGAAGAGGGGCCCAAGAAATTAAGAGTCCCTCAGAGAGAACTATATCTAGTGAAGATG 3780  
 Db 3721 GTTACCGAAGAGGGGCCCAAGAAATTAAGAGTCCCTCAGAGAGAACTATATCTAGTGAAGATG 3780  
 QY 3781 AAGAGCTTCCCTGCTTCCACACACTGTTATTTGTAAGTAAACAATATACCTTCCACT 3840  
 Db 3781 AAGAGCTTCCCTGCTTCCACACACTGTTATTTGTAAGTAAACAATATACCTTCCACT 3840  
 QY 3841 CTACTAGGCAATGACACCGCTTGTCTACCGAGTGTCTGTAAAGAACACAGAGAGAAATTTAT 3900  
 Db 3841 CTACTAGGCAATGACACCGCTTGTCTACCGAGTGTCTGTAAAGAACACAGAGAGAAATTTAT 3900  
 QY 3901 TATCATTTGAAGAAATAGCTTAAATGACTGACGTAACAGGTAATATTTGGCAAGAGCATCTC 3960  
 Db 3901 TATCATTTGAAGAAATAGCTTAAATGACTGACGTAACAGGTAATATTTGGCAAGAGCATCTC 3960  
 QY 3961 AGGAACTACACCTTAGTGAGAGAAACAATGTTCTGCTAGCTTGTCTTCCACAGTGA 4020  
 Db 3961 AGGAACTACACCTTAGTGAGAGAAACAATGTTCTGCTAGCTTGTCTTCCACAGTGA 4020  
 QY 4021 GTGAATTTGGAAGACTTGACGTGCAAAATACAAACCCAGAGATCTTCTTGATTTGGTCTT 4080  
 Db 4021 GTGAATTTGGAAGACTTGACGTGCAAAATACAAACCCAGAGATCTTCTTGATTTGGTCTT 4080  
 QY 4081 CCAAAACAATGAGGCAATCACTCTGAAAGCCAGGGAATTTGGTCTGAGTGCACAAAGATTTGG 4140  
 Db 4081 CCAAAACAATGAGGCAATCACTCTGAAAGCCAGGGAATTTGGTCTGAGTGCACAAAGATTTGG 4140  
 QY 4141 TTTTCAGATGATGAAGAAAGAGAACGGCTTGGAAAGAAAATATATCAAGAGAGCAAGCA 4200  
 Db 4141 TTTTCAGATGATGAAGAAAGAGAACGGCTTGGAAAGAAAATATATCAAGAGAGCAAGCA 4200  
 QY 4201 TGGATTTCAAACTTAGTGAAGCAGCATCTGGGTGTGAGAGTGAAGAACAGCGTCTTGAGG 4260  
 Db 4201 TGGATTTCAAACTTAGTGAAGCAGCATCTGGGTGTGAGAGTGAAGAACAGCGTCTTGAGG 4260



4261 ACTGCTCAGGGCTATCCTCTCAGAGTACATTTTAACCACTCAGACAGAGGATACCATGC 4320  
 4261 ACTGCTCAGGGCTATCCTCTCAGAGTACATTTTAACCACTCAGACAGAGGATACCATGC 4320  
 4321 AACATACCTGATTAAGCTCCAGACAGAAATGGCTGAAGTGTGTAGAACAGC 4380  
 4321 AACATACCTGATTAAGCTCCAGACAGAAATGGCTGAAGTGTGTAGAACAGC 4380  
 4381 ATGGAGCCAGCCCTTCTACAGCTACCTTCCATCATAGAGTCTTCTGCTTGAGG 4440  
 4381 ATGGAGCCAGCCCTTCTACAGCTACCTTCCATCATAGAGTCTTCTGCTTGAGG 4440  
 4441 ACCTGCGAAATCCAGAACAGACATCAGAAAAAGCAGTATTAACTTACAGAAAASTA 4500  
 4441 ACCTGCGAAATCCAGAACAGACATCAGAAAAAGCAGTATTAACTTACAGAAAASTA 4500  
 4501 GTGAATACCTTATTAAGCCAGAAATCCAGAAAGCCCTTCTGCTGACAAAGTTGAGGTCTG 4560  
 4501 GTGAATACCTTATTAAGCCAGAAATCCAGAAAGCCCTTCTGCTGACAAAGTTGAGGTCTG 4560  
 4561 CAGATAGTCTTACAGTAAAAATTAAGAACCCAGAGGTGAAAGGTCATCCCTCTTAAT 4620  
 4561 CAGATAGTCTTACAGTAAAAATTAAGAACCCAGAGGTGAAAGGTCATCCCTCTTAAT 4620  
 4621 GCCCATCTTATGATAGTGTGTACATGCACAGTGTCTGGAGTCTTCAAGATAGAA 4680  
 4621 GCCCATCTTATGATAGTGTGTACATGCACAGTGTCTGGAGTCTTCAAGATAGAA 4680  
 4681 ACTACCCATCTCAAGAGAGCTCATTAAGTGTGTGATGTGAGAGAGCCAGACAGCTGAG 4740  
 4681 ACTACCCATCTCAAGAGAGCTCATTAAGTGTGTGATGTGAGAGAGCCAGACAGCTGAG 4740  
 4741 AGTGTGGGCCACAGATTTGACGGAACATCTTACTTCCAGAGCAGATCTAGAGGAA 4800  
 4741 AGTGTGGGCCACAGATTTGACGGAACATCTTACTTCCAGAGCAGATCTAGAGGAA 4800  
 4801 CCCCTACCTGGAATCTGGAATCAGACCTCTCTGATGAGACCTGGAATCTGATCTG 4860  
 4801 CCCCTACCTGGAATCTGGAATCAGACCTCTCTGATGAGACCTGGAATCTGATCTG 4860  
 4861 AAGACAGAGCCCGAGAGTCACTGCTGTGTGCAACATCACTTCAACCTCTGATGGA 4920  
 4861 AAGACAGAGCCCGAGAGTCACTGCTGTGTGCAACATCACTTCAACCTCTGATGGA 4920  
 4921 AAGTTCCTGATTAAGTGTGCAATCTGCCAGAGTCCAGCTGCTCTATCTACTG 4980  
 4921 AAGTTCCTGATTAAGTGTGCAATCTGCCAGAGTCCAGCTGCTCTATCTACTG 4980  
 4981 ATACTGCTGGGTATTAATGCAATGGAAGAGTGTGAGAGAGGAGAGCAGAAATGACAG 5040  
 4981 ATACTGCTGGGTATTAATGCAATGGAAGAGTGTGAGAGAGGAGAGCAGAAATGACAG 5040  
 5041 CTTCACAGAGAAAGGTCACAAAGAAATGTCATGCTGTGTGCTGAGCCCTGAGCAGAG 5100  
 5041 CTTCACAGAGAAAGGTCACAAAGAAATGTCATGCTGTGTGCTGAGCCCTGAGCAGAG 5100  
 5101 AATTTATGCTGTGTACAGAGTGTGCAAGAAATGTCATGCTGTGTGCTGAGCAGAG 5160  
 5101 AATTTATGCTGTGTACAGAGTGTGCAAGAAATGTCATGCTGTGTGCTGAGCAGAG 5160  
 5161 CTGAGAGAGTACTGATGCTGTGTGAAAGAGATGCTGATGCTGTGTGCTGAGCAGAG 5220  
 5161 CTGAGAGAGTACTGATGCTGTGTGAAAGAGATGCTGATGCTGTGTGCTGAGCAGAG 5220  
 5221 TGAATATTTTCTAGGAATTCGGGAGAGAAATGGGTAGTATGCTATTTCTGGGTGACCC 5280  
 5221 TGAATATTTTCTAGGAATTCGGGAGAGAAATGGGTAGTATGCTATTTCTGGGTGACCC 5280  
 5281 AGTCTATTTAAAGAAAGAAATGCTGAATGAGCATGATTTGAAAGTCCAGAGAGATGTGG 5340  
 5281 AGTCTATTTAAAGAAAGAAATGCTGAATGAGCATGATTTGAAAGTCCAGAGAGATGTGG 5340

5341 TCAATGAGAAACCCAGAGGTCACAAAGCGAGAGAGAAATCCAGAGAGAAAGATCT 5400  
 5341 TCAATGAGAAACCCAGAGGTCACAAAGCGAGAGAGAAATCCAGAGAGAAAGATCT 5400  
 5401 TCAGGGGCTAGAAATCTGTTGCTATGAGGCCCTTCCACCAATGCCCACAGATCACTGG 5460  
 5401 TCAGGGGCTAGAAATCTGTTGCTATGAGGCCCTTCCACCAATGCCCACAGATCACTGG 5460  
 5461 AATGATGATACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5520  
 5461 AATGATGATACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5520  
 5521 GCACAGGTGTCACCCCAATTTGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5580  
 5521 GCACAGGTGTCACCCCAATTTGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5580  
 5581 TCCATGCAATTTGGGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5640  
 5581 TCCATGCAATTTGGGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5640  
 5641 GTTAGCAGCTTACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5700  
 5641 GTTAGCAGCTTACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5700  
 5701 GCCACTACTGA 5711  
 5701 GCCACTACTGA 5711

RESULT 7  
 ARI25601  
 LOCUS ARI25601  
 DEFINITION Sequence 1 from patent US 6177410.  
 ACCESSION ARI25601  
 VERSION ARI25601.1 GI:14111663  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 5712)  
 AUTHORS Holt,D.T., Jensen,R.A., King,M.-C., Steiner,M.S.,  
 Robinson-Benion,C.I. and Thompson,M.E.  
 TITLE Therapeutic methods for prostate cancer  
 JOURNAL Patent: US 6177410-A 1 23 -JAN-2001;  
 FEATURES  
 source Location/Qualifiers  
 1..5712  
 BASE COUNT 1956 a 1099 c 1274 g 1383 t  
 ORIGIN

Query Match 99.9%; Score 5707.8; DB 6; Length 5712;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 5709; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 1 AGCTGCTGAGACCTCTGAGACCCGACAGAGGCTGTGGGTTTCTCAGATTAAGTGGCC 60  
 1 AGCTGCTGAGACCTCTGAGACCCGACAGAGGCTGTGGGTTTCTCAGATTAAGTGGCC 60  
 1 AGCTGCTGAGACCTCTGAGACCCGACAGAGGCTGTGGGTTTCTCAGATTAAGTGGCC 60  
 61 CCTGGCTCAGAGAGCCCTTCAACCTCTGCTGTGGTAAAGTTCATTTGAGAGAGAA 120  
 61 CCTGGCTCAGAGAGCCCTTCAACCTCTGCTGTGGTAAAGTTCATTTGAGAGAGAA 120  
 61 CCTGGCTCAGAGAGCCCTTCAACCTCTGCTGTGGTAAAGTTCATTTGAGAGAGAA 120  
 121 TGGATTATCTGCTCTTGGCGGTGAAGAAGTACAAATGTCTATTAATGCTATGCAAGAAA 180  
 121 TGGATTATCTGCTCTTGGCGGTGAAGAAGTACAAATGTCTATTAATGCTATGCAAGAAA 180  
 121 TGGATTATCTGCTCTTGGCGGTGAAGAAGTACAAATGTCTATTAATGCTATGCAAGAAA 180  
 181 TCTTAGAGTTCACCATCTGCTGATGATCAAGAGACCTGTCTCCCAAGAGTGTGACC 240  
 181 TCTTAGAGTTCACCATCTGCTGATGATCAAGAGACCTGTCTCCCAAGAGTGTGACC 240  
 181 TCTTAGAGTTCACCATCTGCTGATGATCAAGAGACCTGTCTCCCAAGAGTGTGACC 240  
 241 ACATATTTTGAAGAAATTTGATGCTGAAGACTTCTCAACAGAGAGAAAGGGCTTTCACAGT 300  
 241 ACATATTTTGAAGAAATTTGATGCTGAAGACTTCTCAACAGAGAGAAAGGGCTTTCACAGT 300

QY 301 GTCCTTATGTAGAATGATATATACCAAAAGAGCCCTACAGAAGTACAGATTATAGTC 360  
 |||||  
 Db 301 GTCCTTATGTAGAATGATATATACCAAAAGAGCCCTACAGAAGTACAGATTATAGTC 360  
 QY 361 AACCTGTGAAGAGCTATTGAAAATCATTTTGTCTTTTACGCTTGACACAGTTTGGAGT 420  
 |||||  
 Db 361 AACCTGTGAAGAGCTATTGAAAATCATTTTGTCTTTTACGCTTGACACAGTTTGGAGT 420  
 QY 421 ATGCAAAAGCTATATTTTGCAAAAGAAAATACTCTCCGTAACATCTTAAAGATG 480  
 |||||  
 Db 421 ATGCAAAAGCTATATTTTGCAAAAGAAAATACTCTCCGTAACATCTTAAAGATG 480  
 QY 481 AAGTTTCTATCATCCAAAGATGGGCTACAGAAACCGTGCCAAAAGACTTCTACAGATG 540  
 |||||  
 Db 481 AAGTTTCTATCATCCAAAGATGGGCTACAGAAACCGTGCCAAAAGACTTCTACAGATG 540  
 QY 541 AACCCGAAAATCTCTCTTGCGAGAAACCACTGCTCAGTGTCCAACTCTTAACCTTGGAA 600  
 |||||  
 Db 541 AACCCGAAAATCTCTCTTGCGAGAAACCACTGCTCAGTGTCCAACTCTTAACCTTGGAA 600  
 QY 601 CTGTGAGACTCTGAGGACAAAGCAGCGGATACACCTCAAAAGAGCTGTCTACATG 660  
 |||||  
 Db 601 CTGTGAGACTCTGAGGACAAAGCAGCGGATACACCTCAAAAGAGCTGTCTACATG 660  
 QY 661 AATTGGGATCTGATCTTCTGAAAGATACCGTTAATAAGCACTTATTGCAGTGGGAG 720  
 |||||  
 Db 661 AATTGGGATCTGATCTTCTGAAAGATACCGTTAATAAGCACTTATTGCAGTGGGAG 720  
 QY 721 ATCAAGAATTGTTTCAAAATCACCCTCAAGAAACCGGATGAAATCAGTTGGATTCTG 780  
 |||||  
 Db 721 ATCAAGAATTGTTTCAAAATCACCCTCAAGAAACCGGATGAAATCAGTTGGATTCTG 780  
 QY 781 CAAAAAGGCTCTGTGTAATTTTCTGAGACGGATTAACAATACTGAACATCATCAAC 840  
 |||||  
 Db 781 CAAAAAGGCTCTGTGTAATTTTCTGAGACGGATTAACAATACTGAACATCATCAAC 840  
 QY 841 CCAGTAATATGATTTGAACACCACCTGAGAAAGCTGACGTGAGAGGCATCCAGAAAAGT 900  
 |||||  
 Db 841 CCAGTAATATGATTTGAACACCACCTGAGAAAGCTGACGTGAGAGGCATCCAGAAAAGT 900  
 QY 901 ATCAGGGATGTTCTGTTTCAAACTGCAATGTGAGGCACTGTGCACAAATCTCATGCCA 960  
 |||||  
 Db 901 ATCAGGGATGTTCTGTTTCAAACTGCAATGTGAGGCACTGTGCACAAATCTCATGCCA 960  
 QY 961 GCTCATTTACAGCATGAGAACAGCAGTTTATTACTCACTAAAGACAGATGAATGTAGAA 1020  
 |||||  
 Db 961 GCTCATTTACAGCATGAGAACAGCAGTTTATTACTCACTAAAGACAGATGAATGTAGAA 1020  
 QY 1021 AAGCTGAATTTCTGTATTAAGCAAAACAGCTGCTTACAGAGGACCACATACAGAT 1080  
 |||||  
 Db 1021 AAGCTGAATTTCTGTATTAAGCAAAACAGCTGCTTACAGAGGACCACATACAGAT 1080  
 QY 1081 GGGGCTGGAAGTAAGGAAACATGTAATGATAGGGGAGCTCCACAGCAGAAAAGAGTGG 1140  
 |||||  
 Db 1081 GGGGCTGGAAGTAAGGAAACATGTAATGATAGGGGAGCTCCACAGCAGAAAAGAGTGG 1140  
 QY 1141 ATCTGAATGCTGATCCCTGTGTGAGAGAAAAGATGAATAAGCAGAAATGCCATGCT 1200  
 |||||  
 Db 1141 ATCTGAATGCTGATCCCTGTGTGAGAGAAAAGATGAATAAGCAGAAATGCCATGCT 1200  
 QY 1201 CAGAGATTCCTAGACATGTAAGATGTTCTTGGATTAACACTAATATGACGATTCAGA 1260  
 |||||  
 Db 1201 CAGAGATTCCTAGACATGTAAGATGTTCTTGGATTAACACTAATATGACGATTCAGA 1260  
 QY 1261 AAGTTAATGAGTGTTCCTCCAGAAAGTGAAGTGTAGTCTGTGACTCATCAGATG 1320  
 |||||  
 Db 1261 AAGTTAATGAGTGTTCCTCCAGAAAGTGAAGTGTAGTCTGTGACTCATCAGATG 1320  
 QY 1321 GGGAGTCTGAATCAATGCCAAAGTACGTGATGTATTGGACGTTCTTAATGAGGTAGATG 1380  
 |||||  
 Db 1321 GGGAGTCTGAATCAATGCCAAAGTACGTGATGTATTGGACGTTCTTAATGAGGTAGATG 1380  
 QY 1381 AATATCTGTTCTTCAGAGAAAATAGACTTACTGGCCAGATGATCCATGAGGCTTTAA 1440  
 |||||  
 Db 1381 AATATCTGTTCTTCAGAGAAAATAGACTTACTGGCCAGATGATCCATGAGGCTTTAA 1440  
 QY 1441 TATGTAAAGTGAAGAGTTCACCTCCAAATCACTAGAGAGTAATATTGAAGCAAAATAT 1500  
 |||||  
 Db 1441 TATGTAAAGTGAAGAGTTCACCTCCAAATCACTAGAGAGTAATATTGAAGCAAAATAT 1500  
 QY 1501 TTGGGAAAACCTATTCGGAAGAGCAAGCTCCCACTTAAGCCATGATGAAAATC 1560  
 |||||  
 Db 1501 TTGGGAAAACCTATTCGGAAGAGCAAGCTCCCACTTAAGCCATGATGAAAATC 1560  
 QY 1561 TAAATTAAGGACATTTGTTACTAGCCACAGATTAATACAGAGCTGCCCTCACAAATA 1620  
 |||||  
 Db 1561 TAAATTAAGGACATTTGTTACTAGCCACAGATTAATACAGAGCTGCCCTCACAAATA 1620  
 QY 1621 AATTAAAGCTTAAAGAGACCTACATCAGGCTTATCTGAGATTTTATCAGAAAG 1680  
 |||||  
 Db 1621 AATTAAAGCTTAAAGAGACCTACATCAGGCTTATCTGAGATTTTATCAGAAAG 1680  
 QY 1681 CAGATTTGGCAGTTCAAAAAGACTCTGAAATGATAATCAGGGAACCTAACCAAGGAGC 1740  
 |||||  
 Db 1681 CAGATTTGGCAGTTCAAAAAGACTCTGAAATGATAATCAGGGAACCTAACCAAGGAGC 1740  
 QY 1741 AGAATGCTCAAGTATGAATATTACTAATAGTGTGATGATGAGATAAACAAGGTGATT 1800  
 |||||  
 Db 1741 AGAATGCTCAAGTATGAATATTACTAATAGTGTGATGATGAGATAAACAAGGTGATT 1800  
 QY 1801 CTATTCAGATTGAGAAAATCCTTAACCCATAGATCCTCGAAAAGAAATCTGCTTCA 1860  
 |||||  
 Db 1801 CTATTCAGATTGAGAAAATCCTTAACCCATAGATCCTCGAAAAGAAATCTGCTTCA 1860  
 QY 1861 AAACGAAAGCTGAACCTATTAAGCAGAGATTAAGCAATATGAAATCGAATTAATATCC 1920  
 |||||  
 Db 1861 AAACGAAAGCTGAACCTATTAAGCAGAGATTAAGCAATATGAAATCGAATTAATATCC 1920  
 QY 1921 ACAATTTCAAAAGCACCCTAAAGAAATAGGCTGAGAGGAGTCTTCTACAGGATATTTC 1980  
 |||||  
 Db 1921 ACAATTTCAAAAGCACCCTAAAGAAATAGGCTGAGAGGAGTCTTCTACAGGATATTTC 1980  
 QY 1981 ATGCGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2040  
 |||||  
 Db 1981 ATGCGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2040  
 QY 2041 TTGATAGTTGTTCTTACAGTGAAGATTAAGAAAAAAGTAAACCAACCAATGCGAGTCA 2100  
 |||||  
 Db 2041 TTGATAGTTGTTCTTACAGTGAAGATTAAGAAAAAAGTAAACCAACCAATGCGAGTCA 2100  
 QY 2101 GGCACAGCAGAAACCTACATCATGAGAGTGAAGAACCTGAGGACCAAGAGA 2160  
 |||||  
 Db 2101 GGCACAGCAGAAACCTACATCATGAGAGTGAAGAACCTGAGGACCAAGAGA 2160  
 QY 2161 GTAACAGCCAAATGACAGACAGTAAAGAGATGACAGTATCTTCCAGAGCTGA 2220  
 |||||  
 Db 2161 GTAACAGCCAAATGACAGACAGTAAAGAGATGACAGTATCTTCCAGAGCTGA 2220  
 QY 2221 AGTTAACAATGACACCTGTTCTTTTACTAAGTGTCAAAATCCAGTAACTTAAAGAT 2280  
 |||||  
 Db 2221 AGTTAACAATGACACCTGTTCTTTTACTAAGTGTCAAAATCCAGTAACTTAAAGAT 2280  
 QY 2281 TTGTCAATCTTACCTTCCAGAGAGAAAAGAGAGAACTAGAAACAGTTAAAGTGT 2340  
 |||||  
 Db 2281 TTGTCAATCTTACCTTCCAGAGAGAAAAGAGAGAACTAGAAACAGTTAAAGTGT 2340  
 QY 2341 CTAATTAATGCTGAAGCCCCCAAGATCTCATGTTAAGTGAAGAGGTTTGGCAACTG 2400  
 |||||  
 Db 2341 CTAATTAATGCTGAAGCCCCCAAGATCTCATGTTAAGTGAAGAGGTTTGGCAACTG 2400  
 QY 2401 AAAAGATCTAGAGATGACATATTTCTGTTGCTACCTGATGATTAATGACACTGAG 2460  
 |||||  
 Db 2401 AAAAGATCTAGAGATGACATATTTCTGTTGCTACCTGATGATTAATGACACTGAG 2460  
 QY 2461 AAAAGTATCTGTTACTGGAAGTTAGCACTTAGGGAAGCAAAAACAGAACCAATAAAT 2520  
 |||||



Dp	2461	AAAGATCTCTTACTGGAATTTAGCACTCTAGGGAAGGCAAAAAACAGACCAATAAT	2520
Qy	2521	GTGTGAGTCAGTGTGCAAGCATTTGAAAAACCCCAAGGCACTAATTCATGGTTGTTCCAAG	2580
Dp	2521	GTGTGAGTCAGTGTGCAAGCATTTGAAAAACCCCAAGGCACTAATTCATGGTTGTTCCAAG	2580
Qy	2581	ATTAATAGAAATGACACAGAGAGGCTTTAAGTATTCATTGGGACATGAAGTTAACCAAGTC	2640
Dp	2581	ATTAATAGAAATGACACAGAGAGGCTTTAAGTATTCATTGGGACATGAAGTTAACCAAGTC	2640
Qy	2641	GGGAACACAGCATGGAATGGAACAAAGTGAACCTGATGCACTATTTCAGAAATCAT	2700
Dp	2641	GGGAACACAGCATGGAATGGAACAAAGTGAACCTGATGCACTATTTCAGAAATCAT	2700
Qy	2701	TCAGAGTTTCAAAAGCCGACATTTCTCGTTTTCAAATCCAGAAATCCAAAGAG	2760
Dp	2701	TCAGAGTTTCAAAAGCCGACATTTCTCGTTTTCAAATCCAGAAATCCAAAGAG	2760
Qy	2761	AATGTGCAACATTCCTGCCCCACTCTGGGTCCTTAAAGAAACAAAGTCCAAAGTCACCT	2820
Dp	2761	AATGTGCAACATTCCTGCCCCACTCTGGGTCCTTAAAGAAACAAAGTCCAAAGTCACCT	2820
Qy	2821	TTGAAATGGAACAAAGGAAGAAAAATCAGSAAAGATGAGTCTAATATCAAGCTGTAC	2880
Dp	2821	TTGAAATGGAACAAAGGAAGAAAAATCAGSAAAGATGAGTCTAATATCAAGCTGTAC	2880
Qy	2881	AGACAGTTAATATCACTGACGCTTTCCCTGGTGGTCAGAAAGATTAAGCAGTTGATA	2940
Dp	2881	AGACAGTTAATATCACTGACGCTTTCCCTGGTGGTCAGAAAGATTAAGCAGTTGATA	2940
Qy	2941	ATGCCAAATGTAGTATCAAGAGAGGCTCTAGGTTTGTCTATCATCTTCAGTTCCAGAGCA	3000
Dp	2941	ATGCCAAATGTAGTATCAAGAGAGGCTCTAGGTTTGTCTATCATCTTCAGTTCCAGAGCA	3000
Qy	3001	ACGAAACCTGGACCTATTACTCCAAATTAACATGCACTTTTCAAAACCCATATCTGTATAC	3060
Dp	3001	ACGAAACCTGGACCTATTACTCCAAATTAACATGCACTTTTCAAAACCCATATCTGTATAC	3060
Qy	3061	CACACACTTTTCCCATCAAGTCACTTTGTTAAACCTAATGTAAAGAAAAATTCGTAGAGG	3120
Dp	3061	CACACACTTTTCCCATCAAGTCACTTTGTTAAACCTAATGTAAAGAAAAATTCGTAGAGG	3120
Qy	3121	AAAACTTTGAGGAACATTCATGTCCACTGGAAGAGAAATGGAATGAGAACATTTCCAA	3180
Dp	3121	AAAACTTTGAGGAACATTCATGTCCACTGGAAGAGAAATGGAATGAGAACATTTCCAA	3180
Qy	3181	GTACAGTGACACAATTAGCCGTAATAACATTAGAGAAATGTTTTTAAAGAACCCAGCT	3240
Dp	3181	GTACAGTGACACAATTAGCCGTAATAACATTAGAGAAATGTTTTTAAAGAACCCAGCT	3240
Qy	3241	CAACCAATATTAATGAAGTAGCTCCGACTACTAATGAAGTGGGCTCCAGTATTATGAAA	3300
Dp	3241	CAACCAATATTAATGAAGTAGCTCCGACTACTAATGAAGTGGGCTCCAGTATTATGAAA	3300
Qy	3301	TAGGTTCCAGTATGAAAAACATTCAGAGCAACTAGGTAGAAACAGAGGCCCAAAATTGA	3360
Dp	3301	TAGGTTCCAGTATGAAAAACATTCAGAGCAACTAGGTAGAAACAGAGGCCCAAAATTGA	3360
Qy	3361	ATGCTATGCTTAGATTAGGGGTTTTGCAACCTGAGGCTTATTAACAAAGTCTTCTGGAA	3420
Dp	3361	ATGCTATGCTTAGATTAGGGGTTTTGCAACCTGAGGCTTATTAACAAAGTCTTCTGGAA	3420
Qy	3421	GTAATTTGTAACATCTCGAAATTAATAAAGCAAGAAATATGAAGAGTAGTTCAAGCTGTA	3480
Dp	3421	GTAATTTGTAACATCTCGAAATTAATAAAGCAAGAAATATGAAGAGTAGTTCAAGCTGTA	3480
Qy	3481	ATTACAGATTTCTCCCATATCTGATTTTCAGATTAACCTTAGAACACACCTATGGAAGTAGTC	3540
Dp	3481	ATTACAGATTTCTCCCATATCTGATTTTCAGATTAACCTTAGAACACACCTATGGAAGTAGTC	3540
Qy	3541	ATGCACTCTCAGGTTGTTCTGAGACACCTGATGACCTGTAGATGATGTTGAAATTAAGG	3600
Dp	3541	ATGCACTCTCAGGTTGTTCTGAGACACCTGATGACCTGTAGATGATGTTGAAATTAAGG	3600

Dd	3541	ATGCATCTCAGGTTTGTCTCTGACACACCTGATGACCTGTGTAGATGATGCTGAATTAAG	3600
Oy	3601	AAGATCTAGTTTTGCTGGAATAAGACATTTAAAGAAAGTTGCTGTTTTTAAAGAAAACG	3660
Dd	3601	AAGATCTAGTTTTGCTGTAATAATGACATTTAAGAAAGTTGCTGTTTTTAAAGAAAACG	3660
Oy	3661	TCCAGAAAGGAGAGCTTACGACAGAGTCTTAACCTTTTACCCTATACATTTGGCTCAGG	3720
Dd	3661	TCCAGAAAGGAGAGCTTACGACAGAGTCTTAACCTTTTACCCTATACATTTGGCTCAGG	3720
Oy	3721	GTTACCCGAAGGGGGCCAGAAATTAGTGCTCTGGAAGAGACACTATCTAGTGGATG	3780
Dd	3721	GTTACCCGAAGGGGGCCAGAAATTAGTGCTCTCTGAGAGAACTTATCTAGTGGATG	3780
Oy	3781	AAGAGCTTCCCTGCTCCACACCTGTATTGTGTAAGTAACAAATATACCTTCTCAGT	3840
Dd	3781	AAGAGCTTCCCTGCTCCACACCTGTATTGTGTAAGTAACAAATATACCTTCTCAGT	3840
Oy	3841	CTACTAGGCATAGACACGCTTGCTACCGAGTGTCTGTCTAAGAACACAGAGAGAAATTA	3900
Dd	3841	CTACTAGGCATAGACACGCTTGCTACCGAGTGTCTGTCTAAGAACACAGAGAGAAATTA	3900
Oy	3901	TATCATTTAAGAAATATAGCTTAATATGATCTGACTAACCAGTAAATATGTGGCAAGCATCTC	3960
Dd	3901	TATCATTTAAGAAATATAGCTTAATATGATCTGACTAACCAGTAAATATGTGGCAAGCATCTC	3960
Oy	3961	AGGAACATCACCTTAGTGAGAAACAAATTTCTGCTAGCTGTGTTCTTCCACAGTSCA	4020
Dd	3961	AGGAACATCACCTTAGTGAGAAACAAATTTCTGCTAGCTGTGTTCTTCCACAGTSCA	4020
Oy	4021	GTGAATTTGGAAGACTTGACTGCTCAAAATACAAACACCCAGAGTCTTCTTGATTTGGTCTT	4080
Dd	4021	GTGAATTTGGAAGACTTGACTGCTCAAAATACAAACACCCAGAGTCTTCTTGATTTGGTCTT	4080
Oy	4081	CCAAACAATATAGCATCATGCTGTGAANCCAGGAGTTGGTCTGATGACACAAGAAATGG	4140
Dd	4081	CCAAACAATATAGCATCATGCTGTGAANCCAGGAGTTGGTCTGATGACACAAGAAATGG	4140
Oy	4141	TTTCAAGATGATGAAGAAAGAGAAAGGGCTTGGAAGAAATTAATCAAGAACAAACCA	4200
Dd	4141	TTTCAAGATGATGAAGAAAGAGAAAGGGCTTGGAAGAAATTAATCAAGAACAAACCA	4200
Oy	4261	ACTGCTCAGGGCTTACTCTCTCAGAGTGACATTTTAAACCACCTCACAGAGGGATACATGC	4320
Dd	4261	ACTGCTCAGGGCTTACTCTCTCAGAGTGACATTTTAAACCACCTCACAGAGGGATACATGC	4320
Oy	4321	AACATTAACCTATTAAGCTCCAGAGAAATAGGCGAATCAAGAGCTGTGTTAAACACGC	4380
Dd	4321	AACATTAACCTATTAAGCTCCAGAGAAATAGGCGAATCAAGAGCTGTGTTAAACACGC	4380
Oy	4381	ATGGGAGCAGCCCTTCTAACAGCTAACCTTCCATCATTAAGTACCTCTCCCTTGAGG	4440
Dd	4381	ATGGGAGCAGCCCTTCTAACAGCTAACCTTCCATCATTAAGTACCTCTCCCTTGAGG	4440
Oy	4441	ACCTGCGCAAAATCCAGAACAAAGACATTCAGAAAAAGCAGTTTAACTTCAACAGAAAAGTA	4500
Dd	4441	ACCTGCGCAAAATCCAGAACAAAGACATTCAGAAAAAGCAGTTTAACTTCAACAGAAAAGTA	4500
Oy	4501	GTGAATACCCTATTAAGCCAGAAATCCAGAAAGCCCTTCTGCTGACAAAGTTTGAAGTGTCTG	4560
Dd	4501	GTGAATACCCTATTAAGCCAGAAATCCAGAAAGCCCTTCTGCTGACAAAGTTTGAAGTGTCTG	4560
Oy	4561	CAGATAGTCTTACCACTAATAAATTAAGAACCCAGAGTGGAAAGGTCAATCCCTCTTAAT	4620
Dd	4561	CAGATAGTCTTACCACTAATAAATTAAGAACCCAGAGTGGAAAGGTCAATCCCTCTTAAT	4620
Oy	4621	GCCCATCATTAAGATGATAGTGTTGATCATGCAAGTTCCTTGGAAGTCTTCAGAAATAGAA	4680
Dd	4621	GCCCATCATTAAGATGATAGTGTTGATCATGCAAGTTCCTTGGAAGTCTTCAGAAATAGAA	4680

OY 4681 ACTACCATCTCAAGAGAGCTCATTAAGGTTGTTGATGTGAGAGCAACAGCTGAGAG 4740  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 4681 ACTACCATCTCAAGAGAGAGCTCATTAAGGTTGTTGATGTGAGAGCAACAGCTGAGAG 4740  
 OY 4741 AGTCTGGGCCACACAGATTTGACGGAAACATCTTACTTCCAAAGGCAAGATCTAGAGGAA 4800  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 4741 AGTCTGGGCCACACAGATTTGACGGAAACATCTTACTTCCAAAGGCAAGATCTAGAGGAA 4800  
 OY 4801 CCCCTTACCTGGAAATCGGAATCAGCCTCTCTGTGATGAGCCCGGAATCGATCTCTCTG 4860  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 4801 CCCCTTACCTGGAAATCGGAATCAGCCTCTCTGTGATGAGCCCGGAATCGATCTCTCTG 4860  
 OY 4861 AAGACAGAGCCCGCAGAGATCAGTCTGTGTGGCAACATACCATCTTCAACCTCTGATTTGA 4920  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 4861 AAGACAGAGCCCGCAGAGATCAGTCTGTGTGGCAACATACCATCTTCAACCTCTGATTTGA 4920  
 OY 4921 AAGTCCCAATTTGAAAGTTGACAGATCTGCCAGACTCCAGCTGCTGCTCATACTCTG 4980  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 4921 AAGTCCCAATTTGAAAGTTGACAGATCTGCCAGACTCCAGCTGCTGCTCATACTCTG 4980  
 OY 4981 ATACTGCTGGGTAAATGCAATGGAAGAAGTGTGACAGGAGAGAGCCGAATTTGACAG 5040  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 4981 ATACTGCTGGGTAAATGCAATGGAAGAAGTGTGACAGGAGAGAGCCGAATTTGACAG 5040  
 OY 5041 CTTCACAGAGAAAGGCTCAACAAAAGATGTCATGATGATGCTGCTGAGCCCTGACCCAGAG 5100  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 5041 CTTCACAGAGAAAGGCTCAACAAAAGATGTCATGATGATGCTGCTGAGCCCTGACCCAGAG 5100  
 OY 5101 AATTTTGTCTGCTGATCACTGTTGCCAGAAAACACCATCTTAACTTAATCTTAATTA 5160  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 5101 AATTTTGTCTGCTGATCACTGTTGCCAGAAAACACCATCTTAACTTAATCTTAATTA 5160  
 OY 5161 CTGAAGAGACTACTCATGTTGTTATGAAAACAGATGCTGAGTTGTTGTTGATGAGAGAGATG 5220  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 5161 CTGAAGAGACTACTCATGTTGTTATGAAAACAGATGCTGAGTTGTTGTTGATGAGAGAGATG 5220  
 OY 5221 TGAATAATTTTCTAGGAATTTGCGGAGAGAAAATGAGTAACTTATTTGCGGTGACCC 5280  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 5221 TGAATAATTTTCTAGGAATTTGCGGAGAGAAAATGAGTAACTTATTTGCGGTGACCC 5280  
 OY 5281 AGCTCTTTAAAGAAAGAAAATGCTGAAATGAGCATGATTTTGAAGTGAAGAGAGATGAG 5340  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 5281 AGCTCTTTAAAGAAAGAAAATGCTGAAATGAGCATGATTTTGAAGTGAAGAGAGATGAG 5340  
 OY 5341 TCATGGAAGAAACCAACCAAGTCTCAAGAGGAGCAAGAGATCCAGAGAGAGAAAGATCT 5400  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 5341 TCATGGAAGAAACCAACCAAGTCTCAAGAGGAGCAAGAGATCCAGAGAGAGAAAGATCT 5400  
 OY 5401 TCAGGGGGCTAGAAAATCTGTTGCTATGAGGCCCTTACCAACATGCCCCAGATCAATG 5460  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 5401 TCAGGGGGCTAGAAAATCTGTTGCTATGAGGCCCTTACCAACATGCCCCAGATCAATG 5460  
 OY 5461 AATGAGTGTGACAGCTGTGTGCTCTCTGAGAGAGAGCTTCAATCAATCCCTG 5520  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 5461 AATGAGTGTGACAGCTGTGTGCTCTCTGAGAGAGAGCTTCAATCAATCCCTG 5520  
 OY 5521 GCACAGGTGTCCACCAATTTGCTGTGTGACAGCAGATGCTGAGAGAGAGAGATGCT 5580  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 5521 GCACAGGTGTCCACCAATTTGCTGTGTGACAGCAGATGCTGAGAGAGAGAGATGCT 5580  
 OY 5581 TCCATCAATTTGCGCAGATGTGTGAGGACACTGTGTGAGACCCGAGAGTGGTGTGGACA 5640  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 5581 TCCATCAATTTGCGCAGATGTGTGAGGACACTGTGTGAGACCCGAGAGTGGTGTGGACA 5640  
 OY 5641 GTGTAGACTCTTCAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5700  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 5641 GTGTAGACTCTTCAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5700  
 OY 5701 GCCACTACTGA 5711  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 5701 GCCACTACTGA 5711

RESULT 8  
 AR184044  
 LOCUS AR184044 5712 bp DNA linear PAT 20-APR-2002  
 DEFINITION Sequence 47 from patent US 6342483.  
 ACCESSION AR184044  
 VERSION AR184044.1 GI:20228013  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Unknown.  
 REFERENCE  
 1 (bases 1 to 5712)  
 HOLZ, J.T., Jensen, R.A., Page, D.L., Obermiller, P.S.,  
 Robinson-Benion, C.L., and Thompson, M.E.  
 Method for detection and treatment of breast cancer  
 Patent: US 6342483-A 47 29-JUN-2002;  
 FEATURES  
 source  
 1..5712  
 Location/Qualifiers  
 BASE COUNT 1956 a 1099 c 1274 g 1383 t  
 ORIGIN  
 Query Match 99.98; Score 5707.8; DB 6; Length 5712;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 5709; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 OY- 1 AGCTCGCTGAGACTTCTGAGACCCCGCAGCAGGCTGTGGGTTTTCAGATTAACCTGGCC 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 1 AGCTCGCTGAGACTTCTGAGACCCCGCAGCAGGCTGTGGGTTTTCAGATTAACCTGGCC 60  
 OY 61 CCTCGCTCAGAGAGGCTTCAACCTCTGCTGCTGAGTGAAGTTCATTGGAACAGAGAGAA 120  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 61 CCTCGCTCAGAGAGGCTTCAACCTCTGCTGCTGAGTGAAGTTCATTGGAACAGAGAGAA 120  
 OY 121 TGGATTTATGCTGCTTGGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 121 TGGATTTATGCTGCTTGGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180  
 OY 181 TCTTAGAGTGTCCATCTGTCTGAGTGTGATCAAGAACTGCTCCACAAAGTGTGACC 240  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 181 TCTTAGAGTGTCCATCTGTCTGAGTGTGATCAAGAACTGCTCCACAAAGTGTGACC 240  
 OY 241 ACATATTTTGAATTTTTCATGCTGAACTTCTCAACCAAGAGAGAGAGAGAGAGAGAGAG 300  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 241 ACATATTTTGAATTTTTCATGCTGAACTTCTCAACCAAGAGAGAGAGAGAGAGAGAGAG 300  
 OY 301 GTCCCTTATGTAAGATGATATTAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 301 GTCCCTTATGTAAGATGATATTAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360  
 OY 361 AACTTGTGTAAGAGCTATTGAAATCAATTTGTGCTTTTCAAGTGTGACAGAGTTTGGAGT 420  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 361 AACTTGTGTAAGAGCTATTGAAATCAATTTGTGCTTTTCAAGTGTGACAGAGTTTGGAGT 420  
 OY 421 ATGCAAAACAGCTATTAATTTTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 421 ATGCAAAACAGCTATTAATTTTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480  
 OY 481 AAGTTTCTATCATCAAGATGAGGCTACAGAAACCGTGCAGAAAGAGCTTCTACAGAGTG 540  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 481 AAGTTTCTATCATCAAGATGAGGCTACAGAAACCGTGCAGAAAGAGCTTCTACAGAGTG 540  
 OY 541 AACCCGAAATTCCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 541 AACCCGAAATTCCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600  
 OY 601 CTGTGAGAGACTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 601 CTGTGAGAGACTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660  
 OY 661 AATGGAGATCGATCTTCTGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 661 AATGGAGATCGATCTTCTGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720

QY	721	ATCAAGATTTGTTACAAATCACCCCTCAAGAACACGGGATGAATCAAGTTTGATTCG	780
Db	721	ATCAAGAAATTTGTTACAAATCACCCCTCAAGGAACACGGGATGAATCAAGTTTGATTCG	780
QY	781	CAAAAAAGGCTCTGTGTAATTTCTGACAGGATGTACAAATCTGGAACATCATCAAC	840
Db	781	CAAAAAAGGCTCTGTGTAATTTCTGACAGGATGTACAAATCTGGAACATCATCAAC	840
QY	841	CCAGTAATATGATTTTGAACACACACTGAGAACGCTGACACTGAGAGCATCCAGAAAGT	900
Db	841	CCAGTAATATGATTTTGAACACACACTGAGAACGCTGACACTGAGAGCATCCAGAAAGT	900
QY	901	ATCAGGGTAGTTCTGTTTCAAACTTGCAATGTGAGCCATGTGGCACAAATACTCATGCCA	960
Db	901	ATCAGGGTAGTTCTGTTTCAAACTTGCAATGTGAGCCATGTGGCACAAATACTCATGCCA	960
QY	961	GCTCATTTACAGCATGGAACACACAGTTTCTACTACTAAAGCAGATGAATGTGAAA	102
Db	961	GCTCATTTACAGCATGGAACACACAGTTTCTACTACTAAAGCAGATGAATGTGAAA	102
QY	1021	AGGCTGAATTCGTATAATAAAGCAACACCGCTTGCTTACCAAGAGCCACAATACAGAT	108
Db	1021	AGGCTGAATTCGTATAATAAAGCAACACCGCTTGCTTACCAAGAGCCACAATACAGAT	108
QY	1081	GGGCTGGAATTAAGGAACATGTAATGATAGGCGGACTCCCGACACAGAAAAAAGTGAG	114
Db	1081	GGGCTGGAATTAAGGAACATGTAATGATAGGCGGACTCCCGACACAGAAAAAAGTGAG	114
QY	1141	ATCTGAATGCTGATATCCCTGTGTGAGAGAAAAAAGTAATGAAAGCAAACTGCCATGCT	120
Db	1141	ATCTGAATGCTGATATCCCTGTGTGAGAGAAAAAAGTAATGAAAGCAAACTGCCATGCT	120
QY	1201	CAGAGAAATCTTGAAGATACTGAGATGTTCTTGATTAACACTAAATAGCAGCATTCAGA	126
Db	1201	CAGAGAAATCTTGAAGATACTGAGATGTTCTTGATTAACACTAAATAGCAGCATTCAGA	126
QY	1261	AAGTTAATGAGAGGTTTTCCAGACAGATGAAGCTGTAGTCTGTGATGCTCACATGATG	132
Db	1261	AAGTTAATGAGAGGTTTTCCAGACAGATGAAGCTGTAGTCTGTGATGCTCACATGATG	132
QY	1321	GGGAGTCTGAATCAATGCCAAAGTAGCTGATGTATGGACGTTCTTAATGAGGTAGATG	138
Db	1321	GGGAGTCTGAATCAATGCCAAAGTAGCTGATGTATGGACGTTCTTAATGAGGTAGATG	138
QY	1381	AATATTCGTGTTCTCAGAGAAATATAGACTTCTGCGCAGTATCTCTATGAGGCTTAA	144
Db	1381	AATATTCGTGTTCTCAGAGAAATATAGACTTCTGCGCAGTATCTCTATGAGGCTTAA	144
QY	1441	TATGTAAAGTGAAAGAGTTCACCTCAATCAGTAGAGATTAATTTGAAGCAAAATAT	150
Db	1441	TATGTAAAGTGAAAGAGTTCACCTCAATCAGTAGAGATTAATTTGAAGCAAAATAT	150
QY	1501	TTGGGAAAACCTATGCGGAAGAGGCAAGCCTCCCACTTAAGCCATGTATCTGAAATC	156
Db	1501	TTGGGAAAACCTATGCGGAAGAGGCAAGCCTCCCACTTAAGCCATGTATCTGAAATC	156
QY	1561	TAAATATAGAGACATTTGTTACTGAGCCACAGATATATACAGAGCGTCCCTCACAAATA	162
Db	1561	TAAATATAGAGACATTTGTTACTGAGCCACAGATATATACAGAGCGTCCCTCACAAATA	162
QY	1621	AATTTAAACGTTAAAGAGACCTACATCAGGCTTCATCTGAGAGATTTTATCAAGAAAG	168
Db	1621	AATTTAAACGTTAAAGAGACCTACATCAGGCTTCATCTGAGAGATTTTATCAAGAAAG	168
QY	1681	CAGATTTGGAGTTCAAAAAGACTCCTGGAATATATTAATCAGGGAATCAACAAAGGAGC	174
Db	1681	CAGATTTGGAGTTCAAAAAGACTCCTGGAATATATTAATCAGGGAATCAACAAAGGAGC	174
QY	1741	AGAAATGCTCAAGTGATGAATTTACTAATAGTGTGATGAGATTAACAAAGGCTGAT	180
Db	1741	AGAAATGCTCAAGTGATGAATTTACTAATAGTGTGATGAGATTAACAAAGGCTGAT	180

QY	1801	CTATTTCAGATGAGAAAAATCTTAACCCAAATGAACTCCTCGAAAAAGAAATCTGCTTCA	1860
Db	1801	CTATTTCAGAAATGAGAAAAATCTTAACCCAAATGAAATCACTCTGAAAAAAGAAATCTGCTTCA	1860
QY	1861	AAAGCAAAAGCTGAACCTTAATTAAGGACCACTATTAAGCAATATGGAAGTGGAAATATATATTC	1920
Db	1861	AAAGCAAAAGCTGAACCTTAATTAAGGACCACTATTAAGCAATATGGAAGTGGAAATATATATTC	1920
QY	1921	ACAAATTCAAAAAGCACCTAAAAAGATAGCGTAGAGAGAAATCTCTTACAGGCATATTC	1980
Db	1921	ACAAATTCAAAAAGCACCTAAAAAGATAGCGTAGAGAGAAATCTCTTACAGGCATATTC	1980
QY	1981	ATGCGCTTGAACCTAGTAGTCAGTAGAAATCTTAAGCCCACTTAATTTGATAGTAATTTGCAAA	2040
Db	1981	ATGCGCTTGAACCTAGTAGTCAGTAGAAATCTTAAGCCCACTTAATTTGATAGTAATTTGCAAA	2040
QY	2041	TTGATATGTTTCTTAGCAGTGAAGAGATTAAGAAAAAAABAATACACCAATATGCGACTCA	2100
Db	2041	TTGATATGTTTCTTAGCAGTGAAGAGATTAAGAAAAAAABAATACACCAATATGCGACTCA	2100
QY	2101	GGCAGAGAGAAACCTTACACTCATATGGAAGGTAAAGAAACCTGCAACTGAGACCAGAGAA	2160
Db	2101	GGCAGAGAGAAACCTTACACTCATATGGAAGGTAAAGAAACCTGCAACTGAGACCAGAGAA	2160
QY	2161	GTAAACAAGCCAAATGAAACAGACAAGTAAGAAGACATGACAGTACTTCTCCAGAGCTGA	2220
Db	2161	GTAAACAAGCCAAATGAAACAGACAAGTAAGAAGACATGACAGTACTTCTCCAGAGCTGA	2220
QY	2221	AGTTAACCAATATGCACCTGTGTTCTTTACTTAAGTGTCAATACAGTGAACCTTAAGAT	2280
Db	2221	AGTTAACCAATATGCACCTGTGTTCTTTACTTAAGTGTCAATACAGTGAACCTTAAGAT	2280
QY	2281	TTGTCAATCTAGCCCTTCAGAGAGAGAAAAAGAAAGAAATCTGAAACACTTAAAGTGT	2340
Db	2281	TTGTCAATCTAGCCCTTCAGAGAGAGAAAAAGAAAGAAATCTGAAACACTTAAAGTGT	2340
QY	2341	CTAATTAATGCTGAAGAACCCCAAGATCTCATGTTAAGTGAAGGGAAGGGTTTTGCAAACTG	2400
Db	2341	CTAATTAATGCTGAAGAACCCCAAGATCTCATGTTAAGTGAAGGGAAGGGTTTTGCAAACTG	2400
QY	2401	AAAGATCTGTAGAGAGTGAAGATTTTCATTGTGTAACCTGTAAGTATTAAGCACTCAG	2460
Db	2401	AAAGATCTGTAGAGAGTGAAGATTTTCATTGTGTAACCTGTAAGTATTAAGCACTCAG	2460
QY	2461	AAAGATCTGTAGAGAGTGAAGATTTTCATTGTGTAACCTGTAAGTATTAAGCACTCAG	2520
Db	2461	AAAGATCTGTAGAGAGTGAAGATTTTCATTGTGTAACCTGTAAGTATTAAGCACTCAG	2520
QY	2521	GTGTGAGTCACTGTGACACTTTTAAAAACCCCAAGGGGACTAATTCATGAGTGTTCACAAG	2580
Db	2521	GTGTGAGTCACTGTGACACTTTTAAAAACCCCAAGGGGACTAATTCATGAGTGTTCACAAG	2580
QY	2581	ATTAATAGAAATGACACAGAGAGGCTTTTAAGTATTCATTGGACATGAAGTTAACCACAGTC	2640
Db	2581	ATTAATAGAAATGACACAGAGAGGCTTTTAAGTATTCATTGGACATGAAGTTAACCACAGTC	2640
QY	2641	GGGAACAAGCATTTGAATGGAAGAAAGTGAATGATGTCACATATTTTGGAGAAATTCAT	2700
Db	2641	GGGAACAAGCATTTGAATGGAAGAAAGTGAATGATGTCACATATTTTGGAGAAATTCAT	2700
QY	2701	TCAAGGTTTCAAAAGCGCCAGTCAATTTGCTGTTTTCAATTCAGAGAAATGACAGAAAGG	2760
Db	2701	TCAAGGTTTCAAAAGCGCCAGTCAATTTGCTGTTTTCAATTCAGAGAAATGACAGAAAGG	2760
QY	2761	AATGTGCAACATTTCTGTGCCACTCTGTGGTCTTTAAAGAAACAAGTCCAAAAGTCACTT	2820
Db	2761	AATGTGCAACATTTCTGTGCCACTCTGTGGTCTTTAAAGAAACAAGTCCAAAAGTCACTT	2820
QY	2821	TTGAATGTGAACAACAAGGAAGAAATCAAGGAAAGAAATGAGTCTAATATCAAGCGTAC	2880
Db	2821	TTGAATGTGAACAACAAGGAAGAAATCAAGGAAAGAAATGAGTCTAATATCAAGCGTAC	2880
QY	2881	AGACAGTTTAATATCTGACAGGCTTCTGTTGTTGATCGAAAGATATACCCAGTTGAT	2940

```
|||||
Db 2881 AGACAGTTAATATACATGACAGGCTTTCCTGTTGGTCAGAAAGATTAAGCCAGTTGATA 2940
OY 2941 ATGCCAAATAGATATATAAAGGAGGCTAGGTTTGTCTATCATCTCAGTTCCAGAGCA 3000
Db 2941 ATGCCAAATAGATATATAAAGGAGGCTAGGTTTGTCTATCATCTCAGTTCCAGAGCA 3000
OY 3001 ACGAACACTGAGCTATCTACCAATTAACATGAGCTTTTACAAACCCATATGATATAC 3060
Db 3001 ACGAACACTGAGCTATCTACCAATTAACATGAGCTTTTACAAACCCATATGATATAC 3060
OY 3061 CACCACCTTTTCCATCAGCTATTTGTTAAACTAAATGTAAGAAAAATCTGCTAGAG 3120
Db 3061 CACCACCTTTTCCATCAGCTATTTGTTAAACTAAATGTAAGAAAAATCTGCTAGAG 3120
OY 3121 AAAACTTTGAGGAACATTCATGTCACCTGAAAGAAATGGGAAATGAGAAACATTCOA 3180
Db 3121 AAAACTTTGAGGAACATTCATGTCACCTGAAAGAAATGGGAAATGAGAAACATTCOA 3180
OY 3181 GTACAGTGAGCACATTAAGCCGTAAATAACATTAGAGAAAATGTTTTTAAAGACCACT 3240
Db 3181 GTACAGTGAGCACATTAAGCCGTAAATAACATTAGAGAAAATGTTTTTAAAGACCACT 3240
OY 3241 CAAAGCAATATTAAATGAGAGTTCAGTACTAATGAAAGTGGCTCCAGTATTATGAA 3300
Db 3241 CAAAGCAATATTAAATGAGAGTTCAGTACTAATGAAAGTGGCTCCAGTATTATGAA 3300
OY 3301 TAGGTTCCAGTGAATGAACATTCACAGAACTAGAGTAAGAAAGAGGCGCAAAATGGA 3360
Db 3301 TAGGTTCCAGTGAATGAACATTCACAGAACTAGAGTAAGAAAGAGGCGCAAAATGGA 3360
OY 3361 ATGCTATGCTTAGATTAGGGGTTTTGCAACCTGAGTCTATATAACAAAGTTCCTGGA 3420
Db 3361 ATGCTATGCTTAGATTAGGGGTTTTGCAACCTGAGTCTATATAACAAAGTTCCTGGA 3420
OY 3421 GTATTTGTAAAGCATCTGGAATAAAAAGCAAGATATGAAGAACTAGTTCAGACTGTA 3480
Db 3421 GTATTTGTAAAGCATCTGGAATAAAAAGCAAGATATGAAGAACTAGTTCAGACTGTA 3480
OY 3481 ATACAGATTTCTCCATATCTGATTCAGATACTAGAACAGCTATGGAAGTATGTC 3540
Db 3481 ATACAGATTTCTCCATATCTGATTCAGATACTAGAACAGCTATGGAAGTATGTC 3540
OY 3541 ATGCATCTCAGGTTTGTCTGAGACACCTGATGACCTGTTAGATGATGCTGAAATTAAG 3600
Db 3541 ATGCATCTCAGGTTTGTCTGAGACACCTGATGACCTGTTAGATGATGCTGAAATTAAG 3600
OY 3601 AAGATCTACTTTTGTGGAATAATGACATTAAGGAAAGTTCGTGTTTTAGCAAAACG 3660
Db 3601 AAGATCTACTTTTGTGGAATAATGACATTAAGGAAAGTTCGTGTTTTAGCAAAACG 3660
OY 3661 TTCAGAAAGAGAGCTTAGCAGAGAGTCTAGGCCCTTACCCATACATTTGGCTCAG 3720
Db 3661 TTCAGAAAGAGAGCTTAGCAGAGAGTCTAGGCCCTTACCCATACATTTGGCTCAG 3720
OY 3721 GTTACCGAAGAGGAGCCAGAAATTAAGAGTCTCAGAGAGAACTTATCTAGTAGATG 3780
Db 3721 GTTACCGAAGAGGAGCCAGAAATTAAGAGTCTCAGAGAGAACTTATCTAGTAGATG 3780
OY 3781 AAGAGCTTCCCTGCTTCAACACTTGTATTTGTTAAAGTAACAAATATACCTTCTCAG 3840
Db 3781 AAGAGCTTCCCTGCTTCAACACTTGTATTTGTTAAAGTAACAAATATACCTTCTCAG 3840
OY 3841 CTACTGAGGCACTAGCACCCTGCTACCGAGTGTCTGTAAGAAACACAGAGAGAAATTA 3900
Db 3841 CTACTGAGGCACTAGCACCCTGCTACCGAGTGTCTGTAAGAAACACAGAGAGAAATTA 3900
OY 3901 TATCATTTGAAGATAGCTTAATATGATGAGTAACAGTAAATATGGAAGAAGCATCTC 3960
Db 3901 TATCATTTGAAGATAGCTTAATATGATGAGTAACAGTAAATATGGAAGAAGCATCTC 3960
OY 3961 AGGACATCACCTTAGTGAGGAACAAATGTTCTGCTAGCTTTTCTTACAGTGA 4020
Db 3961 AGGACATCACCTTAGTGAGGAACAAATGTTCTGCTAGCTTTTCTTACAGTGA 4020

|||||
Db 3961 AGGAATCATACCTTAGTGAGGAACAAATGTTCTGCTAGCTTTTCTTACAGTGA 4020
OY 4021 GTGAATTTGGAAGACTTACATGCAATATACAAACCCAGATTCCTTCTGATTTGCTT 4080
Db 4021 GTGAATTTGGAAGACTTACATGCAATATACAAACCCAGATTCCTTCTGATTTGCTT 4080
OY 4081 CCAACAAATGAGGATCATGCTGAAAGCCAGGAGTGTCTGATGATGACAAAGATTTG 4140
Db 4081 CCAACAAATGAGGATCATGCTGAAAGCCAGGAGTGTCTGATGATGACAAAGATTTG 4140
OY 4141 TTTTCAGATGATGAAGAAAGAGAAAGGCGTTTGGAAAGAAATTAATCAAGAGAGCAAG 4200
Db 4141 TTTTCAGATGATGAAGAAAGAGAAAGGCGTTTGGAAAGAAATTAATCAAGAGAGCAAG 4200
OY 4201 TTTTCAGATGATGAAGAAAGAGAAAGGCGTTTGGAAAGAAATTAATCAAGAGAGCAAG 4260
Db 4201 TTTTCAGATGATGAAGAAAGAGAAAGGCGTTTGGAAAGAAATTAATCAAGAGAGCAAG 4260
OY 4261 TTTTCAGATGATGAAGAAAGAGAAAGGCGTTTGGAAAGAAATTAATCAAGAGAGCAAG 4320
Db 4261 TTTTCAGATGATGAAGAAAGAGAAAGGCGTTTGGAAAGAAATTAATCAAGAGAGCAAG 4320
OY 4321 AACATTAACCTGATTAAGCTCCAGAGAAATGAGTGAAGAAAGTGTGTTAGAACAGC 4380
Db 4321 AACATTAACCTGATTAAGCTCCAGAGAAATGAGTGAAGAAAGTGTGTTAGAACAGC 4380
OY 4381 ATGGAGGCGAGCTTCTAAGAGTACGCTTCCATCATTAAGTACTGCTTGGCCCTGAG 4440
Db 4381 ATGGAGGCGAGCTTCTAAGAGTACGCTTCCATCATTAAGTACTGCTTGGCCCTGAG 4440
OY 4441 ACCTGCGAATAATCCAGAAACAAAGCAGATCAGAAAAAGCAGTATTAACTTCAGAAAAAG 4500
Db 4441 ACCTGCGAATAATCCAGAAACAAAGCAGATCAGAAAAAGCAGTATTAACTTCAGAAAAAG 4500
OY 4501 GTGAATACCTGATTAAGCTCCAGAGAAATGAGTGAAGAAAGTGTGTTAGAACAGC 4560
Db 4501 GTGAATACCTGATTAAGCTCCAGAGAAATGAGTGAAGAAAGTGTGTTAGAACAGC 4560
OY 4561 CAGATAGCTTACAGTAAATTAAGAAACAGAGAGTGAAGAAAGTGTGTTAGAACAGC 4620
Db 4561 CAGATAGCTTACAGTAAATTAAGAAACAGAGAGTGAAGAAAGTGTGTTAGAACAGC 4620
OY 4621 GGCATCATTTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4680
Db 4621 GGCATCATTTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4680
OY 4681 ACTACCCATCTCAAGAGAGAGCTCATTAAGGTTGTGATGTGAGAGGAGCAACAGCTG 4740
Db 4681 ACTACCCATCTCAAGAGAGAGCTCATTAAGGTTGTGATGTGAGAGGAGCAACAGCTG 4740
OY 4741 AGTCTGGGCCACAGATTTTGCAGAAACATCTTACTGCAAGAGCAAGATCTAGAGGAA 4800
Db 4741 AGTCTGGGCCACAGATTTTGCAGAAACATCTTACTGCAAGAGCAAGATCTAGAGGAA 4800
OY 4801 CCCCTTACCTGGAATCTGGAATGACGCTTCTGATGATGATGATGATGATGATGATG 4860
Db 4801 CCCCTTACCTGGAATCTGGAATGACGCTTCTGATGATGATGATGATGATGATGATG 4860
OY 4861 AAGACAGAGCCAGAGAGTCTGTTGGAACATACATCTTAACTTAACTTAACTTAACT 4920
Db 4861 AAGACAGAGCCAGAGAGTCTGTTGGAACATACATCTTAACTTAACTTAACTTAACT 4920
OY 4921 AAGTCCCAATTAAGATTTGAGAAATCTGCAAGATCTGCAAGATCTGCAAGATCTG 4980
Db 4921 AAGTCCCAATTAAGATTTGAGAAATCTGCAAGATCTGCAAGATCTGCAAGATCTG 4980
OY 4981 ATACTGCTGGTATTAATGCAATGAGAAAGTGTGAGCAAGAGAGCAAGATTTGAGAG 5040
Db 4981 ATACTGCTGGTATTAATGCAATGAGAAAGTGTGAGCAAGAGAGCAAGATTTGAGAG 5040
OY 5041 CTTTACAGAAAGAGGCTTAACAAAGAAATGCTGATGCTGCTGCTGAGCCAGAGAG 5100
Db 5041 CTTTACAGAAAGAGGCTTAACAAAGAAATGCTGATGCTGCTGCTGAGCCAGAGAG 5100
```

QY 5101 AATTATGCTGCTGACAGTTTGGCAGAAAACACCATCATTAACTAATCTAATTA 5160  
 DB 5101 AATTATGCTGCTGACAGTTTGGCAGAAAACACCATCATTAACTAATCTAATTA 5160  
 QY 5161 CTGAAGAGACTACTGATGTTGTTATGAAAACAGATGCTGATTTGTGTGAACGAGAC 5220  
 DB 5161 CTGAAGAGACTACTGATGTTGTTATGAAAACAGATGCTGATTTGTGTGAACGAGAC 5220  
 QY 5221 TGAATATTTTCTAGCAATTTGGGAGAGAAATGGGTAGTTAGCTATTTCGTGGTGACC 5280  
 DB 5221 TGAATATTTTCTAGCAATTTGGGAGAGAAATGGGTAGTTAGCTATTTCGTGGTGACC 5280  
 QY 5281 AGCTATTAAAGAAAATGCTGAATGACATGATTTTGAAGTCAAGAGAGATGAG 5340  
 DB 5281 AGCTATTAAAGAAAATGCTGAATGACATGATTTTGAAGTCAAGAGAGATGAG 5340  
 QY 5341 TCAATGAGAAAGAACCAAGGTCCTCAAGAGAGAAATCCAGAGACAGAAAGATCT 5400  
 DB 5341 TCAATGAGAAAGAACCAAGGTCCTCAAGAGAGAAATCCAGAGACAGAAAGATCT 5400  
 QY 5401 TCAGGGGCTGAAATCTGTTGCTATGGGCTTGCACCAACATGCCCACAGATCAACTGG 5460  
 DB 5401 TCAGGGGCTGAAATCTGTTGCTATGGGCTTGCACCAACATGCCCACAGATCAACTGG 5460  
 QY 5461 AATGATGATGACAGCTGTGTGCTGCTGTGTGAAGAGCTTTCATCATTCACCCCTTG 5520  
 DB 5461 AATGATGATGACAGCTGTGTGCTGCTGTGTGAAGAGCTTTCATCATTCACCCCTTG 5520  
 QY 5521 GCACAGGTGTCACCCCAATGTTGTTGTGTCAGCAGCATGCTCGAGACAGAAATGCT 5580  
 DB 5521 GCACAGGTGTCACCCCAATGTTGTTGTGTCAGCAGCATGCTCGAGACAGAAATGCT 5580  
 QY 5581 TCCATGCAATGGGAGATGCTGTGAGGCACTGTGTGATCCGAGAGTGGGTGGACA 5640  
 DB 5581 TCCATGCAATGGGAGATGCTGTGAGGCACTGTGTGATCCGAGAGTGGGTGGACA 5640  
 QY 5641 GTGTAGCACTACAGTGCAGAGAGCTGAGACCTACTGATACCCAGATCCCCACA 5700  
 DB 5641 GTGTAGCACTACAGTGCAGAGAGCTGAGACCTACTGATACCCAGATCCCCACA 5700  
 QY 5701 GCCACTACTGA 5711  
 DB 5701 GCCACTACTGA 5711  
 RESULT 9  
 ARO04673 5914 bp DNA linear PAT 04-DEC-1998  
 LOCUS ARO04673  
 DEFINITION Sequence 1 from patent US 5747282.  
 ACCESSION ARO04673  
 VERSION ARO04673.1 GI:3965552  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 5914)  
 AUTHORS Skolnick M.H., Goldgar D.E., Miki Y., Swenson J., Kanb A.,  
 Harshman K.D., Shattuck-Eidens D.M., Tavtiglian S.V., Wiseman R.W.  
 TITLE 170-linked breast and ovarian cancer susceptibility gene  
 JOURNAL Patent: US 5747282-A 1 05-MAY-1998;  
 FEATURES Location/Qualifiers  
 source 1..5914  
 ORIGIN /organism="Unknown"  
 BASE COUNT 2006 a 1156 c 1316 g 1436 t  
 Query Match 99.9% Score 5707.8; DB 6; Length 5914;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 5709; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 AGCTGCTGAGACTTCTGAGACCCGACAGAGCTGTGGGTTTCTCAGATTAAGTGGCC 60

DB 1 AGCTGCTGAGACTTCTGAGACCCGACAGAGCTGTGGGTTTCTCAGATTAAGTGGCC 60  
 QY 61 CCTGGCTCAGAGAGGCTTACCTCTCTGTCTGGTAAAGTTCAATGGACGAAGAAA 120  
 DB 61 CCTGGCTCAGAGAGGCTTACCTCTCTGTCTGGTAAAGTTCAATGGACGAAGAAA 120  
 QY 121 TGGATTTATCTCTCTGCGCTTGAAGAGTCAAAATGCTATTAATGCTATGCAAAA 180  
 DB 121 TGGATTTATCTCTCTGCGCTTGAAGAGTCAAAATGCTATTAATGCTATGCAAAA 180  
 QY 181 TCTTGAAGTCCATCTCTCTGAGATTGATCAAGAACTCTCTCCAAAGTGTGACC 240  
 DB 181 TCTTGAAGTCCATCTCTCTGAGATTGATCAAGAACTCTCTCCAAAGTGTGACC 240  
 QY 241 ACATATTTTGCAGAAATTTGATGCTGAAACTTCTCAACGAGAAAGAGGCTTACAGT 300  
 DB 241 ACATATTTTGCAGAAATTTGATGCTGAAACTTCTCAACGAGAAAGAGGCTTACAGT 300  
 QY 301 GTCCTTTATGTAGAAATGATATTAACCAAGAGGCTCAAGAAAGTACAGATTTAGTC 360  
 DB 301 GTCCTTTATGTAGAAATGATATTAACCAAGAGGCTCAAGAAAGTACAGATTTAGTC 360  
 QY 361 AACTTTGTAAGAGCTATTTGAAATCATTTGCTTTTCAAGCTTGACACAGTTTGAAGT 420  
 DB 361 AACTTTGTAAGAGCTATTTGAAATCATTTGCTTTTCAAGCTTGACACAGTTTGAAGT 420  
 QY 421 ATGCAAAACGCTATATTTTGCAGAAAGAAATTAAGTCTCTCAACATCTAAAGATG 480  
 DB 421 ATGCAAAACGCTATATTTTGCAGAAAGAAATTAAGTCTCTCAACATCTAAAGATG 480  
 QY 481 AAGTTTATCATCTCAAAAGTATGGGCTACAGAAACGCTGCAAAAGCTCTACAGAGT 540  
 DB 481 AAGTTTATCATCTCAAAAGTATGGGCTACAGAAACGCTGCAAAAGCTCTACAGAGT 540  
 QY 541 AACCCGAAATCTCTCTTCCAGAGAAACAGCTCAGTGTCCAACTCTTAACCTTGGAA 600  
 DB 541 AACCCGAAATCTCTCTTCCAGAGAAACAGCTCAGTGTCCAACTCTTAACCTTGGAA 600  
 QY 601 CTGTAGAGACTGTGAGGACAAAGAGCGGATACAACTCTCAAAAGCGTGTCTACATTG 660  
 DB 601 CTGTAGAGACTGTGAGGACAAAGAGCGGATACAACTCTCAAAAGCGTGTCTACATTG 660  
 QY 661 AATGGAGATGATCTCTCTGAGATACGTTAATAAGCAATTAATGAGTGGGAG 720  
 DB 661 AATGGAGATGATCTCTCTGAGATACGTTAATAAGCAATTAATGAGTGGGAG 720  
 QY 721 ATCAAGAAATGTTTACAAATCAACCCCTCAAGGAACGAGGATGAATCAGTTTGAATCTG 780  
 DB 721 ATCAAGAAATGTTTACAAATCAACCCCTCAAGGAACGAGGATGAATCAGTTTGAATCTG 780  
 QY 781 CAAAAAAGCGTCTGTGATTTTCTGAGAGCGATGAACAAATCTCAACATCTCAAC 840  
 DB 781 CAAAAAAGCGTCTGTGATTTTCTGAGAGCGATGAACAAATCTCAACATCTCAAC 840  
 QY 841 CCAGTAATTAATGATTTGAACACCACTGAGAGCGGAGCGAGAGGATCCAGAAAAGT 900  
 DB 841 CCAGTAATTAATGATTTGAACACCACTGAGAGCGGAGCGAGAGGATCCAGAAAAGT 900  
 QY 901 ATCAGAGTAGTCTCTTCAAACTTGCATGTGAGAGCCATGTTGGCAAAATCTCATGCA 960  
 DB 901 ATCAGAGTAGTCTCTTCAAACTTGCATGTGAGAGCCATGTTGGCAAAATCTCATGCA 960  
 QY 961 GCTCATTTACAGATGAGAGACGAGTTTATTAAGTCAATTAAGAGCAATTAATGAGAAA 1020  
 DB 961 GCTCATTTACAGATGAGAGACGAGTTTATTAAGTCAATTAAGAGCAATTAATGAGAAA 1020  
 QY 1021 AGGCTGAATTTCTGTATTAAGCAAGACCGGCTTGAAGAGGACCAATATACAGAT 1080  
 DB 1021 AGGCTGAATTTCTGTATTAAGCAAGACCGGCTTGAAGAGGACCAATATACAGAT 1080  
 QY 1081 GGGCTGGAAGTAAAGAAACATGTATGATAGCGGACTCCAGACACAGAAAAAGTAG 1140  
 DB 1081 GGGCTGGAAGTAAAGAAACATGTATGATAGCGGACTCCAGACACAGAAAAAGTAG 1140

Db 1081 GGGCTGGAAGTAAAGAAACATGTATAGCGGACCTCCAGCAGCAAGAAAAAGTAG 1140  
QY 1141 ATCTGAATGCTGATCCCTGTGTGAGAGAAAAAATGGAATAGCAGAACTGCCATGCT 1200  
Db 1141 ATCTGAATGCTGATCCCTGTGTGAGAGAAAAAATGGAATAGCAGAACTGCCATGCT 1200  
QY 1201 CAGAGAACTCTAGAGATCTGAGAGATGTTCTTGGATTAACACTTAATAGCAGCATTCAGA 1260  
Db 1201 CAGAGAACTCTAGAGATCTGAGAGATGTTCTTGGATTAACACTTAATAGCAGCATTCAGA 1260  
QY 1261 AAGTAATGAGTGGTTTCCAGAGTGTATGTAAGTGTATGTTGATGATACACTGATGATG 1320  
Db 1261 AAGTAATGAGTGGTTTCCAGAGTGTATGTAAGTGTATGTTGATGATACACTGATGATG 1320  
QY 1321 GGGAGTCTGAATCAAAATGCCAAATAGCTGATGATTGGAGCTTCTAATAGAGTAGATG 1380  
Db 1321 GGGAGTCTGAATCAAAATGCCAAATAGCTGATGATTGGAGCTTCTAATAGAGTAGATG 1380  
QY 1381 AATATTCTGTTCTTCAGAGAAATAGACTTACTGGCCAGTGTCTTCATGAGGCTTTAA 1440  
Db 1381 AATATTCTGTTCTTCAGAGAAATAGACTTACTGGCCAGTGTCTTCATGAGGCTTTAA 1440  
QY 1441 TATGTAAAGTGAAGAGTTCACCTCCAAATCAGTAGAGATTAATTTGAACACAAATAT 1500  
Db 1441 TATGTAAAGTGAAGAGTTCACCTCCAAATCAGTAGAGATTAATTTGAACACAAATAT 1500  
QY 1501 TTGGGAAAACCTATCGGAAAGAGGCAAGCCCTCCCACTTAAGCCATTAAGTGAATATC 1560  
Db 1501 TTGGGAAAACCTATCGGAAAGAGGCAAGCCCTCCCACTTAAGCCATTAAGTGAATATC 1560  
QY 1561 TAAATTATAGACATTTTGTGTACTGAGCCACAGATTAATTAACAAGAGCTCCCTCAAAATA 1620  
Db 1561 TAAATTATAGAGCATTTTGTGTACTGAGCCACAGATTAATTAACAAGAGCTCCCTCAAAATA 1620  
QY 1621 AATTTAAAGCTTAAAGAGACCTATCATCAGGCTTCATCCAGAGATTTTTCAGAAAG 1680  
Db 1621 AATTTAAAGCTTAAAGAGACCTATCATCAGGCTTCATCCAGAGATTTTTCAGAAAG 1680  
QY 1681 CAGATTTGGCAGTTCAAAGAGCTCTGAATATGATTAATCAGGGAACCTAACCAAGGAGC 1740  
Db 1681 CAGATTTGGCAGTTCAAAGAGCTCTGAATATGATTAATCAGGGAACCTAACCAAGGAGC 1740  
QY 1741 AGAATGTCAGAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1800  
Db 1741 AGAATGTCAGAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1800  
QY 1801 CTATTCAGATGAGAAATCTTAACCAATAGAAATACATCGAAAGAAATCTGCTTTCA 1860  
Db 1801 CTATTCAGATGAGAAATCTTAACCAATAGAAATACATCGAAAGAAATCTGCTTTCA 1860  
QY 1861 AAACGAAAGCTGAACCTATTAAGCAGCAGTATTAAGCAATATGGAATCGAATTAATATCC 1920  
Db 1861 AAACGAAAGCTGAACCTATTAAGCAGCAGTATTAAGCAATATGGAATCGAATTAATATCC 1920  
QY 1921 ACAATTCAAAAGCAGCTTAAGAAATAGGCTGAGGAGGAAGTCTTCTACACAGGATATTC 1980  
Db 1921 ACAATTCAAAAGCAGCTTAAGAAATAGGCTGAGGAGGAAGTCTTCTACACAGGATATTC 1980  
QY 1981 ATGCGCTTGAACCTAGTAGTAGTAAGAAATCTAAGCCACCTAATTTAGTACTGAATTTGAAA 2040  
Db 1981 ATGCGCTTGAACCTAGTAGTAGTAAGAAATCTAAGCCACCTAATTTAGTACTGAATTTGAAA 2040  
QY 2041 TTGATAGTGTCTAGCAGTGAAGAGATTAAGAAAAAAGTACAAACCAATGCGACATCA 2100  
Db 2041 TTGATAGTGTCTAGCAGTGAAGAGATTAAGAAAAAAGTACAAACCAATGCGACATCA 2100  
QY 2101 GGCACAGCAGAAACCTTAACAACCTATGGAAGTAAAGAACTGCAACTGAGGAGCCAAAGAA 2160  
Db 2101 GGCACAGCAGAAACCTTAACAACCTATGGAAGTAAAGAACTGCAACTGAGGAGCCAAAGAA 2160  
QY 2161 GTACAGAGCAGAAACCTTAACAACCTATGGAAGTAAAGAACTGCAACTGAGGAGCCAAAGAA 2220  
Db 2161 GTACAGAGCAGAAACCTTAACAACCTATGGAAGTAAAGAACTGCAACTGAGGAGCCAAAGAA 2220

QY 2221 AGTTAAACAATGCACTTGTTCTTTTACTAAGTGTCAATATACAGTAGAACTTAAGAAAT 2280  
Db 2221 AGTTAAACAATGCACTTGTTCTTTTACTAAGTGTCAATATACAGTAGAACTTAAGAAAT 2280  
QY 2281 TTGTCAATCTGACCTTCCAGAGAGAAAAAAGAGAAACCTAAGAAAGTTAAAGTGT 2340  
Db 2281 TTGTCAATCTGACCTTCCAGAGAGAAAAAAGAGAAACCTAAGAAAGTTAAAGTGT 2340  
QY 2341 CTAATATGCTGAAGACCCCAAGATCTCAATGTTAACTGAGAAAGGCTTTTCCAAACTG 2400  
Db 2341 CTAATATGCTGAAGACCCCAAGATCTCAATGTTAACTGAGAAAGGCTTTTCCAAACTG 2400  
QY 2401 AAAGATCTGAGAGTAGCAGTATTTCAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2460  
Db 2401 AAAGATCTGAGAGTAGCAGTATTTCAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2460  
QY 2461 AAAGATCTGAGAGTAGCAGTATTTCAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2520  
Db 2461 AAAGATCTGAGAGTAGCAGTATTTCAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2520  
QY 2521 GTGTGAGTCACTGTCAGCAGTATTTGAAACCCCAAGGAGCTAATTCATGTTGTTCCAAAG 2580  
Db 2521 GTGTGAGTCACTGTCAGCAGTATTTGAAACCCCAAGGAGCTAATTCATGTTGTTCCAAAG 2580  
QY 2581 ATATATAGAAATGACACAGAAAGCTTTAAGTATCCATTGGAGACATGAAGTTAACACAGTC 2640  
Db 2581 ATATATAGAAATGACACAGAAAGCTTTAAGTATCCATTGGAGACATGAAGTTAACACAGTC 2640  
QY 2641 GGGAAACAGCATTAAGAAATGGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2700  
Db 2641 GGGAAACAGCATTAAGAAATGGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2700  
QY 2701 TCAAGGTTTCAACGCGCAGTATTTGCTGTTTCAATTCAGAGAAATGCGAAGAGAG 2760  
Db 2701 TCAAGGTTTCAACGCGCAGTATTTGCTGTTTCAATTCAGAGAAATGCGAAGAGAG 2760  
QY 2761 AATGTGCAACATTTCTGCCCCACTGCTGCTGCTTAAAGAAACAAAGTCCAAAGTCACTT 2820  
Db 2761 AATGTGCAACATTTCTGCCCCACTGCTGCTGCTTAAAGAAACAAAGTCCAAAGTCACTT 2820  
QY 2821 TTGAATGTGAACAAAGAGAAATGCAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2880  
Db 2821 TTGAATGTGAACAAAGAGAAATGCAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2880  
QY 2881 AGACAGTATTAATGACGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2940  
Db 2881 AGACAGTATTAATGACGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2940  
QY 2941 ATGCAATATTAATGACGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3000  
Db 2941 ATGCAATATTAATGACGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3000  
QY 3001 ACAGAACTGGAACCTATTAACCAATTAACCAATTAACCAATTAACCAATTAACCAATTAACCAAT 3060  
Db 3001 ACAGAACTGGAACCTATTAACCAATTAACCAATTAACCAATTAACCAATTAACCAATTAACCAAT 3060  
QY 3061 CACCACTTTTCCCATTAACCAATTAACCAATTAACCAATTAACCAATTAACCAATTAACCAAT 3120  
Db 3061 CACCACTTTTCCCATTAACCAATTAACCAATTAACCAATTAACCAATTAACCAATTAACCAAT 3120  
QY 3121 AAAAATTTGAGGAACATTTCAATGTCACCTGAAAGAGAAATGGAATGGAACATTTCCAA 3180  
Db 3121 AAAAATTTGAGGAACATTTCAATGTCACCTGAAAGAGAAATGGAATGGAACATTTCCAA 3180  
QY 3181 GTACAGTGAACCAATTAACCAATTAACCAATTAACCAATTAACCAATTAACCAATTAACCAAT 3240  
Db 3181 GTACAGTGAACCAATTAACCAATTAACCAATTAACCAATTAACCAATTAACCAATTAACCAAT 3240  
QY 3241 CAAGCAATTAATTAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3300  
Db 3241 CAAGCAATTAATTAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3300



QY	3301	TAGGTTCCAGATGTAACAACTTCACGACGAACCTAGTGAAGAAACAGAGGCCCAATTGA	3360
Db	3301		
QY	3361	ATGCTATGCTAGATTAGGGGTTTTGCAACTGAGGCTTATTAACAAAGTCTTCTGGAA	3420
Db	3361		
QY	3421	GTAATTTGAACTTCCTGAATAAAAAAGCAAGATATGAAGATGATTCAGACTGTTA	3480
Db	3421		
QY	3481	ATACAGATTTCTTCATATCTGATTTCACTTAAGAACAGCTTAGGGAAAGTACTC	3540
Db	3481		
QY	3541	ATGCACTTCAGGTTGTTCTGACACCTGATGACCTGTTGATGATGATGATTAAGG	3600
Db	3541		
QY	3601	AAGATCTAGTTTGTCTGAAATATGACATTTAAGAAAGTTCTGCTTTTGAACAAAGC	3660
Db	3601		
QY	3661	TTCCAGAAAGAGAGCTTAGCAGAGTCTTACCCATACACATTGGCTCAGG	3720
Db	3661		
QY	3721	GTTACCCGAGAGGGGCCAAGAAATTTAGAGTCTCTCAGAAAGAACTTATCTAGTAGAGT	3780
Db	3721		
QY	3781	AAGAGCTTCCCTGCTCCACACTTGTATTTGGTAAAGTAAACATATACCTTCAGT	3840
Db	3781		
QY	3841	CTACTAGGCATAGACCGTGTCTACCGAGTGTCTCTAAGACACAGAGAGAAATTTAT	3900
Db	3841		
QY	3901	TATCATTTGAATAATAGCTTAATGACTGACAGTAACCCAGGTAAATTTGGCAAAAGCATCTC	3960
Db	3901		
QY	3961	AGGAACATCACCTTAGTAGAGAAACAAATGTCTGACTGTTTCTTCAACAGTCA	4020
Db	3961		
QY	4021	GTTGAATTTGGAAGACTTGACTGCAAAATACAAACCCAGAGATCTTTCTTGATTTGTTCTT	4080
Db	4021		
QY	4081	CCAAACAAATAGAGCATCAGTCTGAAAGCCAGGAGTTGGTCTGAGTACAGAAATTTGG	4140
Db	4081		
QY	4141	TTTCAGATGATGAAGAAAGAGAGGAGGCTTGAAGAAATTAATCAAGAAAGAGAAAGCA	4200
Db	4141		
QY	4201	TGGATTTCAAACTTAGGTGAAGCAGCATCTGGGTGAGAGTGAAGAAACAGCGTCTCTGAG	4260
Db	4201		
QY	4261	ACTGCTCAGGGCTATCTCTCAGAGTACATTTTAACCTCAGCAGAGGAGATACCATGTC	4320
Db	4261		
QY	4321	AACATTAACCTTAAGAGTCCAGACGAGAAATGGCGAATAGAGCTGTGTTACAAACGC	4380
Db	4321		
QY	4381	ATGGAGCCAGCCTTCTTACAGCTAACCTTCCATCATTAAGTACCTTCTCTCCCTTGAGG	4440
Db	4381		
QY	4441	ACCTGCAAAATCCGAAACAAAGCAGATCCAGAAAAAGCAGTATTACTTCCAGAAAAAGTA	4500
Db	4441		
QY	4501	GTAATTAACCTTATAGCCAGAAATCCAGAAAGGCTTCTGCTGACAAAGTTTGAGGTGCTG	4560
Db	4501		
QY	4561	CAGATAGTTCTTACAGTAAAAATTAAGAACCCAGAGTGAAGAGGTCACTCCCTTCTTAAT	4620
Db	4561		
QY	4621	GGCCCATTTAGATGATGAGTGGTGAATGACAGATGCTCTGGGAGGTCCTCAGATATGAA	4680
Db	4621		
QY	4681	ACTACCCATCTCAAGAGAGCTCATTAAGTTGTTGATGTGAGAGACACAGCTGGAG	4740
Db	4681		
QY	4741	AGTCTGGCCACACAGATTTGACGGAACATCTTACTTCCAGAGGCAAGATCTAGAGGAA	4800
Db	4741		
QY	4801	CCCCCTTACCTGGAATCTGGAATTCAGCCCTTCTCTGATGACCCCTGAATCTGATCTCTG	4860
Db	4801		
QY	4861	AAGACAGAGCCCGAGAGTCAAGCTGTGTGGCAACATACCATCTTCAACCTCTGCAATGA	4920
Db	4861		
QY	4921	AAGTCCCAATTTGAAGTTGACAGATCTGCCAGAGTCCAGCTGTGCTCATCTACTG	4980
Db	4921		
QY	4981	ATATCTGGGGTAAATGCAATGGAACAAAGTGGACAGGAGAACCCGAATTTGACAG	5040
Db	4981		
QY	5041	CTTCCAGAGAAAGGGTCAACAAAGAAATGTCATGAGTGGTGTGCGGCTCAGCCAGAG	5100
Db	5041		
QY	5101	AATTTATGCTGCTGTACAGATTTGCCAGAAACCCACATCACTTTAACTAATCTAATTA	5160
Db	5101		
QY	5161	CTGAAAGACTACTCATGTTGTTATGAAGAACAGATGCTGAGTTTGTGTGAACGGGACAC	5220
Db	5161		
QY	5221	TGAATATATTTCTGGAATTTGCGGAGAGAAATGGTAGTAGCTATTTCTGGGTGACCC	5280
Db	5221		
QY	5281	AGTCTATTTAAAGAAAGAAATGCTGAATGATGATTTTGAAGTCAAGAGAGATGTGG	5340
Db	5281		
QY	5341	TCAATGGAAGAACCCCAAGGTTCAAGAGGAGCAAGAGATCCAGAGACAGAAAGATCT	5400
Db	5341		
QY	5401	TCAGGGGCTTAGAATCTGTTGCTAATGGGCTTCAACCAATAGCCCAAGATCAACTGG	5460
Db	5401		
QY	5461	AATGATGATGACAGCTGTGCTGCTCTGATGATGAAGAGCTTTCATCAATCACTCTG	5520
Db	5461		

Db 5461 AATGATGTAGACAGCTGTGTGCTCTGTGTGTGAAGAGACCTTTCATTCATCCACCTTG 5520  
Oy 5521 GCACAGGTGTCAACCCCAATTTGTGTGTGCAGACAGATGCTTGACAGAGACATGGCT 5580  
Db 5521 GCACAGGTGTCAACCCCAATTTGTGTGTGCAGACAGATGCTTGACAGAGACATGGCT 5580  
Oy 5581 TCCATGCAATTTGGGACAGATGTGTGAGGACCTGTGTGTGACCCGAGAGTGGTGTGACA 5640  
Db 5581 TCCATGCAATTTGGGACAGATGTGTGAGGACCTGTGTGTGACCCGAGAGTGGTGTGACA 5640  
Oy 5641 GTGTAGACACTCTACCAATGCGCAGAGAGCTGTGACACCTTACCTATATCCCAAGTCCCA 5700  
Db 5641 GTGTAGACACTCTACCAATGCGCAGAGAGCTGTGACACCTTACCTATATCCCAAGTCCCA 5700  
Oy 5701 GCCACTACTGA 5711  
Db 5701 GCCACTACTGA 5711

RESULT 10  
AR008159 5914 bp DNA Linear PAT 04-DEC-1998  
LOCUS DEFINITION AR008159 Sequence 1 from patent US 5753441.  
ACCESSION AR008159 GI:3967268  
VERSION AR008159.1  
KEYWORDS Unknown.  
SOURCE Unknown.  
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 5914)  
AUTHORS Skolnick,M.H., Goldgar,D.E., Miki,Y., Swenson,J., Kamb,A.,  
Harshtman,K.D., Shattuck-Eidens,D.M., Tavtiglian,S.V., Wiseman,R.W.  
and Futreal,P.Andrew.  
170-linked breast and ovarian cancer susceptibility gene  
Patent: US 5753441-A 1 19-MAY-1998;  
TITLE JOURNAL Location/Qualifiers  
FEATURES  
source 1. 5914  
BASE COUNT 2006 a 1156 c 1316 g 1436 t  
ORIGIN

Query Match 99.98; Score 5707.8; DB 6; Length 5914;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 5709; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 AGCTCGCTGAGACTTTCCTGACCCCGCAGAGCTGTGGGTTCTCAGATACTGGGC 60  
Db 1 AGCTCGCTGAGACTTTCCTGACCCCGCAGAGCTGTGGGTTCTCAGATACTGGGC 60  
Oy 61 CCTGCGCTCAGAGAGGCTTACCCCTGCTCTGAGTAAGTTCAATTGGAACAGAAAGAA 120  
Db 61 CCTGCGCTCAGAGAGGCTTACCCCTGCTCTGAGTAAGTTCAATTGGAACAGAAAGAA 120  
Oy 121 TGGATTATCTGCTCTTCGCGTTGAGAGATACAAAATGTCATTATATGTCAGAAAA 180  
Db 121 TGGATTATCTGCTCTTCGCGTTGAGAGATACAAAATGTCATTATATGTCAGAAAA 180  
Oy 181 TCTTAAGTGTCCCATCTGCTGTGAGTGTATCAAGAACTGTCTCCACAAGGTGAGC 240  
Db 181 TCTTAAGTGTCCCATCTGCTGTGAGTGTATCAAGAACTGTCTCCACAAGGTGAGC 240  
Oy 241 ACATATATTTGCAAAATTTGCGATGCTGAACTTCTCAACAGAGAAAGGCGCTTACAGT 300  
Db 241 ACATATATTTGCAAAATTTGCGATGCTGAACTTCTCAACAGAGAAAGGCGCTTACAGT 300  
Oy 301 GTTCCTTATGTAGATGATATATCAACAAAGAGAGCTTACAGAAAGTATGATGTC 360  
Db 301 GTTCCTTATGTAGATGATATATCAACAAAGAGAGCTTACAGAAAGTATGATGTC 360  
Oy 361 AACTGTGTGAAGAGATTTGAATCATTTGCTTTCAGTTGACAGAGTTTGGAGT 420  
Db 361 AACTGTGTGAAGAGATTTGAATCATTTGCTTTCAGTTGACAGAGTTTGGAGT 420

Oy 421 ATGCAACAGCTATATTTTGCACAAAAGAAATAATCTCCTGTAACATCTAAAGAG 480  
Db 421 ATGCAACAGCTATATTTTGCACAAAAGAAATAATCTCCTGTAACATCTAAAGAG 480  
Oy 481 AAGTTTCTATCATCAAAAGATGGGCTACAGAAACCGTGCACAAAGACTTCTACAGATG 540  
Db 481 AAGTTTCTATCATCAAAAGATGGGCTACAGAAACCGTGCACAAAGACTTCTACAGATG 540  
Oy 541 AACCCGAAATCCCTCTGAGAGAAACAGTCTCAGTGTCCAGTCTCTAACCTTGGAA 600  
Db 541 AACCCGAAATCCCTCTGAGAGAAACAGTCTCAGTGTCCAGTCTCTAACCTTGGAA 600  
Oy 601 CTGTAGAACTCTGTAGAGACAAAGCAGCGGATATACACCTCAAAAGAGCTGTCTACATTTG 660  
Db 601 CTGTAGAACTCTGTAGAGACAAAGCAGCGGATATACACCTCAAAAGAGCTGTCTACATTTG 660  
Oy 661 AATTGGATCTGATCTTCTTGTGAGATATCCGTTAATAAGCAACTTATTGACGTGTGGAG 720  
Db 661 AATTGGATCTGATCTTCTTGTGAGATATCCGTTAATAAGCAACTTATTGACGTGTGGAG 720  
Oy 721 ATCAAGAAATGTATCAAAATCACCCTTCAGAGACAGAGGATGAATCAGTTGGATTCTG 780  
Db 721 ATCAAGAAATGTATCAAAATCACCCTTCAGAGACAGAGGATGAATCAGTTGGATTCTG 780  
Oy 781 CAAAAAGAGCTGCTTGTGAATTTTCTGAGACGAGATGTAAACAAATACTGAACATCATCAAC 840  
Db 781 CAAAAAGAGCTGCTTGTGAATTTTCTGAGACGAGATGTAAACAAATACTGAACATCATCAAC 840  
Oy 841 CCACTAATATGATTTTAAACACACACAGAGCTGAGCTGAGAGGATCCAGAAAGT 900  
Db 841 CCACTAATATGATTTTAAACACACACAGAGCTGAGCTGAGAGGATCCAGAAAGT 900  
Oy 901 ATCAGGTAGTCTGCTTCAAACTGTGATGTGAGGATGTGACCAAAATACTCATGCCA 960  
Db 901 ATCAGGTAGTCTGCTTCAAACTGTGATGTGAGGATGTGACCAAAATACTCATGCCA 960  
Oy 961 GCTCATTTACAGCATGAGAACAGCAGTTTATTACTCATTAAGACAGAAATGATGAAA 1020  
Db 961 GCTCATTTACAGCATGAGAACAGCAGTTTATTACTCATTAAGACAGAAATGATGAAA 1020  
Oy 1021 AGGCTGAATCTGTGAATTAAGCAACAGCTGCTTGTAGCAAGAGCCCAATTAACAGAT 1080  
Db 1021 AGGCTGAATCTGTGAATTAAGCAACAGCTGCTTGTAGCAAGAGCCCAATTAACAGAT 1080  
Oy 1081 GGGCTGGAAGTGAAGAAACATGTATGATGTGGGAGCTCCACACAGAAAAAAGGTAG 1140  
Db 1081 GGGCTGGAAGTGAAGAAACATGTATGATGTGGGAGCTCCACACAGAAAAAAGGTAG 1140  
Oy 1141 ATCTGAATGCTGATCCCTGTGTGAGAGAAAGATGGAATGAGCAAGAACTGCCATGCT 1200  
Db 1141 ATCTGAATGCTGATCCCTGTGTGAGAGAAAGATGGAATGAGCAAGAACTGCCATGCT 1200  
Oy 1201 CAGAGAACTCTAAGATACAGAAAGTGTCTTGGATTAACATTAATACACATTCACA 1260  
Db 1201 CAGAGAACTCTAAGATACAGAAAGTGTCTTGGATTAACATTAATACACATTCACA 1260  
Oy 1261 AAGTTAATGAGTGTCTTCCAGAAAGTGAAGTATGATGTTAGCTGTGACATCATGATG 1320  
Db 1261 AAGTTAATGAGTGTCTTCCAGAAAGTGAAGTATGATGTTAGCTGTGACATCATGATG 1320  
Oy 1321 GGGAGTCTGAATCAAAATGCCAAAGTACCTGATGTATGAGACGTTCTAAATGAGTAGATG 1380  
Db 1321 GGGAGTCTGAATCAAAATGCCAAAGTACCTGATGTATGAGACGTTCTAAATGAGTAGATG 1380  
Oy 1381 AATATCTGCTGTCTTCCAGAAAGTATGACTTACTGGCCAGTATCTCTATGAGGCTTTAA 1440  
Db 1381 AATATCTGCTGTCTTCCAGAAAGTATGACTTACTGGCCAGTATCTCTATGAGGCTTTAA 1440  
Oy 1441 TATGTAAAGTGAAGAGTCACTCCAAATCAGTACAGATATATGTAAGCAAAATAT 1500  
Db 1441 TATGTAAAGTGAAGAGTCACTCCAAATCAGTACAGATATATGTAAGCAAAATAT 1500  
Oy 1501 TTGGGAAACCTATCGGAAGAGGCAAGCTCCCACTTAAGCCATGTAACGTAAATTC 1560



```
|||||
Db 1501 TTGGGAAAAACCTATGGAAGGAGGAGCTCCCAACTTAAGCCATGTAACTGAAAAATC 1560
QY 1561 TAATATATAGSACATTTGTTACTGAGCCACAGATATACAAAGAGCTCCCTCACAAAAA 1620
Db 1561 TAATATATAGSACATTTGTTACTGAGCCACAGATATATACAAAGAGCTCCCTCACAAAAA 1620
QY 1621 AATTAAGCGTAAAGAGAGACCTACATCAGGCTTCATCTGAGATTTTATCAAGAAAG 1680
Db 1621 AATTAAGCGTAAAGAGAGACCTACATCAGGCTTCATCTGAGATTTTATCAAGAAAG 1680
QY 1681 CAGATTTGGCAGCTCAAAAAGACTCCTGAAATGATTAATCAGGAACTAAACCAAGGAGC 1740
Db 1681 CAGATTTGGCAGCTCAAAAAGACTCCTGAAATGATTAATCAGGAACTAAACCAAGGAGC 1740
QY 1741 AATAAGTCATAGTATGATATTTACTAATAGTGCATACAGATTAATAACAAAGGTGAT 1800
Db 1741 AATAAGTCATAGTATGATATTTACTAATAGTGCATACAGATTAATAACAAAGGTGAT 1800
QY 1801 CTATTTCAGATGAGAAAAATCCTAACCCATAGAAATCACTCGAAAAAGAAATCTGCTTCA 1860
Db 1801 CTATTTCAGATGAGAAAAATCCTAACCCATAGAAATCACTCGAAAAAGAAATCTGCTTCA 1860
QY 1861 AAACGAAAGCTGACCTATTAAGCAGAGATTAAGCAATATGGAACCTCGAATTAATATCC 1920
Db 1861 AAACGAAAGCTGACCTATTAAGCAGAGATTAAGCAATATGGAACCTCGAATTAATATCC 1920
QY 1921 ACAATTTCAAAAGCAGCTTAAAGAAATAGGCTGAGGAGGAGTCTTCCAGGACATATTC 1980
Db 1921 ACAATTTCAAAAGCAGCTTAAAGAAATAGGCTGAGGAGGAGTCTTCCAGGACATATTC 1980
QY 1981 ATGCGCTTGAACTAGTACGTAGTAAGTAAAGCCACCTAATTTGACTGATTTGCANA 2040
Db 1981 ATGCGCTTGAACTAGTACGTAGTAAGTAAAGCCACCTAATTTGACTGATTTGCANA 2040
QY 2041 TTGATAGTTGTTCTAGCAGTGAAGATTAAGAAAAAAAGTACAAACCAATGCCAGTCA 2100
Db 2041 TTGATAGTTGTTCTAGCAGTGAAGATTAAGAAAAAAAGTACAAACCAATGCCAGTCA 2100
QY 2101 GGCACAGCAGAAACCTACAACTCATGGAAGTAAAGAACTCGCACTGGAGCCAGAAAGA 2160
Db 2101 GGCACAGCAGAAACCTACAACTCATGGAAGTAAAGAACTCGCACTGGAGCCAGAAAGA 2160
QY 2161 GTAAACAGCCAAATGAACAGACAGATTAAGAAAGATGACAGTACTTTCCAGAGCTGA 2220
Db 2161 GTAAACAGCCAAATGAACAGACAGATTAAGAAAGATGACAGTACTTTCCAGAGCTGA 2220
QY 2221 AGTTAAACAATGCACCTGCTTTCTTTACTAATGTTCAAAATACCAAGTAAAGAAAT 2280
Db 2221 AGTTAAACAATGCACCTGCTTTCTTTACTAATGTTCAAAATACCAAGTAAAGAAAT 2280
QY 2281 TTGTAATGCTAGCTTCCAAAGAGAAAGAAAGAGAACTAAGAAACAGTAAAGTGT 2340
Db 2281 TTGTAATGCTAGCTTCCAAAGAGAAAGAAAGAGAACTAAGAAACAGTAAAGTGT 2340
QY 2341 CTAAATATGCTGAAGACCCCAAGATCTCATGTTAAGTGAGAAAGGGTTTTCACAACTG 2400
Db 2341 CTAAATATGCTGAAGACCCCAAGATCTCATGTTAAGTGAGAAAGGGTTTTCACAACTG 2400
QY 2401 AAAAGTCTGTAGAGAGTACAGTATTTCACTTGTACTGTTACTGATTAATGCACTCAGG 2460
Db 2401 AAAAGTCTGTAGAGAGTACAGTATTTCACTTGTACTGTTACTGATTAATGCACTCAGG 2460
QY 2461 AAAGATATCTGTTACTGGAAGTTAGCAGCTTGGGAAAGGCAAAAACAGAACCAATTAAT 2520
Db 2461 AAAGATATCTGTTACTGGAAGTTAGCAGCTTGGGAAAGGCAAAAACAGAACCAATTAAT 2520
QY 2521 GTGTAGTCAAGTGTGACAGATTTGAAAAACCCCAAGGAGTAAATTCATGTTTCCAAAG 2580
Db 2521 GTGTAGTCAAGTGTGACAGATTTGAAAAACCCCAAGGAGTAAATTCATGTTTCCAAAG 2580
QY 2581 ATAATAGAAATGACACAGAGGCTTTAAGTATCCATTGGGACATGAAGTTAACACAGTGC 2640
|||||

Db 2581 ATAATAGAAATGACACAGAGGCTTTAAGTATCCATTGGGACATGAAGTTAACACAGTGC 2640
QY 2641 GGGAAACAAGCATAGAAATGSAAGAAAGATGATCTGATCTAGATTTGGCAATATCAT 2700
Db 2641 GGGAAACAAGCATAGAAATGSAAGAAAGATGATCTGATCTAGATTTGGCAATATCAT 2700
QY 2701 TCAGGTTTCAAGGCGCAGTATTTGCTCTCTTTTCAATATCAGGAAATGCGAAGAGG 2760
Db 2701 TCAGGTTTCAAGGCGCAGTATTTGCTCTCTTTTCAATATCAGGAAATGCGAAGAGG 2760
QY 2761 AATGTGCAACATTTCTGCTCCACTGCTGGTCTTAAAGAACAAAGTCCAAAGTCACTT 2820
Db 2761 AATGTGCAACATTTCTGCTCCACTGCTGGTCTTAAAGAACAAAGTCCAAAGTCACTT 2820
QY 2821 TTGAATGTGAACAAAAGSAGAAATCAAGSAGAAAGATGATCTAATCAAGCCGTATC 2880
Db 2821 TTGAATGTGAACAAAAGSAGAAATCAAGSAGAAAGATGATCTAATCAAGCCGTATC 2880
QY 2881 AGACAGTTAATATCACTGAGGCTTTCCTGTGTGTGTCAGAAAGATAAGCCAGTTGATA 2940
Db 2881 AGACAGTTAATATCACTGAGGCTTTCCTGTGTGTGTCAGAAAGATAAGCCAGTTGATA 2940
QY 2941 ATGCCAATGTATGATCAAGAGAGGCTTAGTCTTTGCTATATCTCAGTTACAGGCA 3000
Db 2941 ATGCCAATGTATGATCAAGAGAGGCTTAGTCTTTGCTATATCTCAGTTACAGGCA 3000
QY 3001 ACGAAACCTGACATCTATCCAAATTAACATGAGACTTTTACAAAACCATATCTGATAC 3060
Db 3001 ACGAAACCTGACATCTATCCAAATTAACATGAGACTTTTACAAAACCATATCTGATAC 3060
QY 3061 CACCACTTTTCCATCAAGATCTTTGTTAAACTAAATGTAAGAAAAATCTGCTAGAGG 3120
Db 3061 CACCACTTTTCCATCAAGATCTTTGTTAAACTAAATGTAAGAAAAATCTGCTAGAGG 3120
QY 3121 AAAAATTGAGAACATTTAATGTCACCTGAAAGAGAAATGCGAAATGCAACATTTCCA 3180
Db 3121 AAAAATTGAGAACATTTAATGTCACCTGAAAGAGAAATGCGAAATGCAACATTTCCA 3180
QY 3181 GTACAGTGAAGCAATTTAGCCGTAATTAACATTAAGAAATGTTTAAAGAAAGCAGCT 3240
Db 3181 GTACAGTGAAGCAATTTAGCCGTAATTAACATTAAGAAATGTTTAAAGAAAGCAGCT 3240
QY 3241 CAAGCAATATTAATGAAGTGTCCAGTACTAATGAAAGTGGGCTCCAGTATTAATGAAA 3300
Db 3241 CAAGCAATATTAATGAAGTGTCCAGTACTAATGAAAGTGGGCTCCAGTATTAATGAAA 3300
QY 3301 TAGGTTCCAGTATGAAACATTTCAAGCAGAACTAGTGAAGAAAGAGGCGCAAAATTTGA 3360
Db 3301 TAGGTTCCAGTATGAAACATTTCAAGCAGAACTAGTGAAGAAAGAGGCGCAAAATTTGA 3360
QY 3361 ATGCTATGCTTATGATTAGGGGTTTTGCAACCTGAGGCTATTAACAAAGTCTCCCGGAA 3420
Db 3361 ATGCTATGCTTATGATTAGGGGTTTTGCAACCTGAGGCTATTAACAAAGTCTCCCGGAA 3420
QY 3421 GTAATTTGAAGCATCTGAAATTAAGCAAGAAATATGMAAGATAGTTACAGCTGTGA 3480
Db 3421 GTAATTTGAAGCATCTGAAATTAAGCAAGAAATATGMAAGATAGTTACAGCTGTGA 3480
QY 3481 ATACAGATTTCTCTCCATATCTGATTTTCAAGATTAAGTGAACAGCTATGGAAGTATGC 3540
Db 3481 ATACAGATTTCTCTCCATATCTGATTTTCAAGATTAAGTGAACAGCTATGGAAGTATGC 3540
QY 3541 ATGCAATCAGGTTTGTCTGTGAGACACCTGATGACCTGTTAGATATATGTTGAATTAAGG 3600
Db 3541 ATGCAATCAGGTTTGTCTGTGAGACACCTGATGACCTGTTAGATATATGTTGAATTAAGG 3600
QY 3601 AAGATAGTATGTTTCTGTAATAATGACATTAAGAAAGTCTGCTTTTACCAAAAGCG 3660
Db 3601 AAGATAGTATGTTTCTGTAATAATGACATTAAGAAAGTCTGCTGTTTACCAAAAGCG 3660
QY 3661 TCCAGAAAAGAGAGGCTTAGCAGAGAGTCCAGCCCTTACACCATACATTTGGCTCAGG 3720
Db 3661 TCCAGAAAAGAGAGGCTTAGCAGAGAGTCCAGCCCTTACACCATACATTTGGCTCAGG 3720
```

OY	3721	GTACCGAAGAGGGCCAGAAATTTAGAGTCTCAGAAAGAACTTATCTAGTGAGATG	3780
Db	3721	GTATCCGAAGAGGGGCCAGAAATTTAGAGTCTCAGAAAGAACTTATCTAGTGAGATG	3780
OY	3781	AAGAGCTTCCCTGCTTCCAGACATTTGTTATTTGGTAAAGTAAACATATACCTTCTCAGT	3840
Db	3781	AAGAGCTTCCCTGCTTCCAGACATTTGTTATTTGGTAAAGTAAACATATACCTTCTCAGT	3840
OY	3841	CTACTAGCATATGACCGCTTCTACCGAGTGTCTGTCTAAGAACACAGAGAGAAATTTAT	3900
Db	3841	CTACTAGCATATGACCGCTTCTACCGAGTGTCTGTCTAAGAACACAGAGAGAAATTTAT	3900
OY	3901	TATCATTTGAAGAAATAGCTTTAAATGATCTGCAATACACAGTAAATATTTGCAAAAGCATCTC	3960
Db	3901	TATCATTTGAAGAAATAGCTTTAAATGATCTGCAATACACAGTAAATATTTGCAAAAGCATCTC	3960
OY	3961	AGGAACATCAOCTTAGTGAAGAGAAACAAATTTCTGCTAGTGTGTTTCTTCCACAGTCA	4020
Db	3961	AGGAACATCAOCTTAGTGAAGAGAAACAAATTTCTGCTAGTGTGTTTCTTCCACAGTCA	4020
OY	4021	GTGAATTTGAAGAGATCTGACATCAATACAAACACCCAGATCCTTTCTGATTTGCTCTT	4080
Db	4021	GTGAATTTGAAGAGATCTGACATCAATACAAACACCCAGATCCTTTCTGATTTGCTCTT	4080
OY	4081	CCAAACAAATGAGGCATCAGTCTGAAGCCAGGAGTTGTTCTGAGTGACAAGAAATGG	4140
Db	4081	CCAAACAAATGAGGCATCAGTCTGAAGCCAGGAGTTGTTCTGAGTGACAAGAAATGG	4140
OY	4141	TTTCAGATGATGAAGAAAGAGGAGCGGCTTGAAGAAATATATCAAGAAAGCAACCA	4200
Db	4141	TTTCAGATGATGAAGAAAGAGGAGCGGCTTGAAGAAATATATCAAGAAAGCAACCA	4200
OY	4201	TGGAATTTAACTTAGTGAAGCAGCATCTGGGTGAGAGTGAAGAAACAGCGCTCTGTAAG	4260
Db	4201	TGGAATTTAACTTAGTGAAGCAGCATCTGGGTGAGAGTGAAGAAACAGCGCTCTGTAAG	4260
OY	4261	ACTGCTCAGGGCTATCTCTCAGAGTGAACATTTTAAACCATCTCAGCAGAGGATACCATGC	4320
Db	4261	ACTGCTCAGGGCTATCTCTCAGAGTGAACATTTTAAACCATCTCAGCAGAGGATACCATGC	4320
OY	4321	AACATTAACCTGATTAAGCTCCAGCAGAAATGGCTGAACCTGAAGCTGTGTTAAGAACGC	4380
Db	4321	AACATTAACCTGATTAAGCTCCAGCAGAAATGGCTGAACCTGAAGCTGTGTTAAGAACGC	4380
OY	4381	ATGGAGCCAGCCCTTCTAAGAGCTAACCTTCCATCATATAGTACTCTTCTGCCCTTGAGG	4440
Db	4381	ATGGAGCCAGCCCTTCTAAGAGCTAACCTTCCATCATATAGTACTCTTCTGCCCTTGAGG	4440
OY	4441	ACCTGCGAAATCCAGAAACAAAGCAGATCAGAAAAAGCAGTATTAACCTTCAAGAAAACTA	4500
Db	4441	ACCTGCGAAATCCAGAAACAAAGCAGATCAGAAAAAGCAGTATTAACCTTCAAGAAAACTA	4500
OY	4501	GTGAATACCTATTAAGCAGAAATCCAGAAAGCCCTTCTGCGACAAAGTTTAAAGGCTGTG	4560
Db	4501	GTGAATACCTATTAAGCAGAAATCCAGAAAGCCCTTCTGCGACAAAGTTTAAAGGCTGTG	4560
OY	4561	CAGATAGTTCTACAGTAAATTAAGAACACAGAGTGAAGAGTCAATCCCTTCTAAAT	4620
Db	4561	CAGATAGTTCTACAGTAAATTAAGAACACAGAGTGAAGAGTCAATCCCTTCTAAAT	4620
OY	4621	GCCCATCATTTAGATGATAGTGTGATGCAACAGTTGCTGTGGAGTCTTGAGAAATAGAA	4680
Db	4621	GCCCATCATTTAGATGATAGTGTGATGCAACAGTTGCTGTGGAGTCTTGAGAAATAGAA	4680
OY	4681	ACTAACCATCTCAAGAGAGAGCTCATTAAGTTGTTGATGTGAGAGACACAGCTGGAAG	4740
Db	4681	ACTAACCATCTCAAGAGAGAGCTCATTAAGTTGTTGATGTGAGAGACACAGCTGGAAG	4740
OY	4741	AGTCTGGGCCACAGATTTGACGGAACATCTTACTTGCCCAAGCAGATCTAGAGGAA	4800
Db	4741	AGTCTGGGCCACAGATTTGACGGAACATCTTACTTGCCCAAGCAGATCTAGAGGAA	4800
OY	4801	CCCTTACCTGGAATCTGGAATCAOCCCTTCTCTGTATGATGACCCGTAATCTGATCCTTCTG	4860
Db	4801	CCCTTACCTGGAATCTGGAATCAOCCCTTCTCTGTATGATGACCCGTAATCTGATCCTTCTG	4860
OY	4861	AAGACAGAGCCCCAGAGTCACTGTGTGGCAACATACATCTTCAACTCTGATTTGA	4920
Db	4861	AAGACAGAGCCCCAGAGTCACTGTGTGGCAACATACATCTTCAACTCTGATTTGA	4920
OY	4921	AAGTCCCAATTTGAAGTTGACAGATCTGCCAGAGTCCAGCTGCTGCTATACCTCTG	4980
Db	4921	AAGTCCCAATTTGAAGTTGACAGATCTGCCAGAGTCCAGCTGCTGCTATACCTCTG	4980
OY	4981	ATACTGCTGGTATTAATGCAATGAAGAAGTGTGACAGAGGAGACCCAGAAATGACAG	5040
Db	4981	ATACTGCTGGTATTAATGCAATGAAGAAGTGTGACAGAGGAGACCCAGAAATGACAG	5040
OY	5041	CTTCAACAGAAAGGCTCAACAAAGAAATGTCATGAGTGTGTGGCTTACCCCAAGAG	5100
Db	5041	CTTCAACAGAAAGGCTCAACAAAGAAATGTCATGAGTGTGTGGCTTACCCCAAGAG	5100
OY	5101	AATTTATGCTGCTGACAGTTTCCAGAAACACACATCATTAACTAAATTAATTA	5160
Db	5101	AATTTATGCTGCTGACAGTTTCCAGAAACACACATCATTAACTAAATTAATTA	5160
OY	5161	CTGAAGAGACTATCATGTTGTTATGAAGACAGATGCTGAGTTGTGTGAACGACAC	5220
Db	5161	CTGAAGAGACTATCATGTTGTTATGAAGACAGATGCTGAGTTGTGTGAACGACAC	5220
OY	5221	TGAATATTTTCTAGGAATTTGGGAGAGAAATGGTGTAGTTAGTATTTCTGGGTGACC	5280
Db	5221	TGAATATTTTCTAGGAATTTGGGAGAGAAATGGTGTAGTTAGTATTTCTGGGTGACC	5280
OY	5281	AGCTATTTAAAGAAAGAAATGCTGATGATGAGATTTTGAAGTACAGAGATGTG	5340
Db	5281	AGCTATTTAAAGAAAGAAATGCTGATGATGAGATTTTGAAGTACAGAGATGTG	5340
OY	5341	TCATGGAAGAAACACCAAGGTCCTGGAAGGAGAGAGATCCAGACAGAAAGATCT	5400
Db	5341	TCATGGAAGAAACACCAAGGTCCTGGAAGGAGAGAGATCCAGACAGAAAGATCT	5400
OY	5401	TCAGGGGCTGGAATCTGTTGATGAGGCTTACCAACATGCTCCAGATCAACTG	5460
Db	5401	TCAGGGGCTGGAATCTGTTGATGAGGCTTACCAACATGCTCCAGATCAACTG	5460
OY	5461	AATGATGATGACACTGTGTGCTGTGTGTGAGAGAGCTTTCATATTCACCTTG	5520
Db	5461	AATGATGATGACACTGTGTGCTGTGTGTGAGAGAGCTTTCATATTCACCTTG	5520
OY	5521	GCACAGGTGTCACCCCAATTTGTGTGTGACACCAATGCTTGACAGAGAGACAATGCT	5580
Db	5521	GCACAGGTGTCACCCCAATTTGTGTGTGACACCAATGCTTGACAGAGAGACAATGCT	5580
OY	5581	TCATGCAATTTGGGAGATGTGTGAGGACCTGTGTGTGACCCAGAGTGGCTTGGACA	5640
Db	5581	TCATGCAATTTGGGAGATGTGTGAGGACCTGTGTGTGACCCAGAGTGGCTTGGACA	5640
OY	5641	GTGAGACACTTACAGTGCAGAGAGCTGGAACATCTGATATACCCCAATCCACACA	5700
Db	5641	GTGAGACACTTACAGTGCAGAGAGCTGGAACATCTGATATACCCCAATCCACACA	5700
OY	5701	GCCACTACTGA 5711	
Db	5701	GCCACTACTGA 5711	
RESULT 11			
ARI36942 5914 bp DNA linear PAT 16-JUN-2001			
LOCUS ARI36942			
DEFINITION Sequence 1 from patent US 6162897.			
ACCESSION ARI36942			
VERSION ARI36942.1 GI:14478192			
KEYWORDS			
SOURCE unknown.			

ORGANISM Unknown.  
 Unclassified.  
 1 (bases 1 to 5914)  
 REFERENCE Skolnick,M.H., Goldgar,D.E., Miki,Y., Swenson,J., Kamb,A.,  
 AUTHORS Harshman,K.D., Shattuck-Eidens,D.M., Tavtigian,S.V., Wiseman,R.W.,  
 and Futreal,P.,Andrew.  
 TITLE 17q-linked breast and ovarian cancer susceptibility gene  
 JOURNAL Patent: US 6162897-A 1 19-DEC-2000;  
 FEATURES Location/Qualifiers  
 source 1..5914  
 /organism="unknown"  
 BASE COUNT 2006 a 1156 c 1316 g 1436 t  
 ORIGIN  
 Query Match 99.9%; Score 5707.8; DB 6; Length 5914;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 5709; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 ACCTGGCTGAGACCTCTCGAGACCCGACACAGGCTGTGGGTTCTCGAGATTAACCTGGGC 60  
 1 AGCTGCTGAGACTTCTCGAGACCCGACACAGGCTGTGGGTTCTCGAGATTAACCTGGGC 60  
 1 AGCTGCTGAGACTTCTCGAGACCCGACACAGGCTGTGGGTTCTCGAGATTAACCTGGGC 60  
 61 CCTGGCTCAGAGAGCCCTTACCCCTCTGCTGGGTAAGTTCATTGGACAGAAAGAA 120  
 61 CCTGGCTCAGAGAGCCCTTACCCCTCTGCTGGGTAAGTTCATTGGACAGAAAGAA 120  
 121 TGGATTATCTGCTCTTGGCGTTGAAGAATGTCATTATGCTATGCAAGAA 180  
 121 TGGATTATCTGCTCTTGGCGTTGAAGAATGTCATTATGCTATGCAAGAA 180  
 121 TGGATTATCTGCTCTTGGCGTTGAAGAATGTCATTATGCTATGCAAGAA 180  
 181 TCTTGAAGTGTCCATCTGTCTGGAGTTGATCAAGAAACCTGTCCCAAGAAAGTGTACC 240  
 181 TCTTGAAGTGTCCATCTGTCTGGAGTTGATCAAGAAACCTGTCCCAAGAAAGTGTACC 240  
 241 ACATATTTTGGCAATTTTGGCATGCTGAAACTCTCAACCAAGAAAGAGGCTTACACAGT 300  
 241 ACATATTTTGGCAATTTTGGCATGCTGAAACTCTCAACCAAGAAAGAGGCTTACACAGT 300  
 301 GTCCCTTATGTAGAATGATATACCAAAAGAGGCTTACCAAGAAAGTACAGATTTAGTC 360  
 301 GTCCCTTATGTAGAATGATATACCAAAAGAGGCTTACCAAGAAAGTACAGATTTAGTC 360  
 361 AACCTGTGAGAGCTATTTGAAATCATTTGCTTTTCAAGCTTACACAGGTTGGAGT 420  
 361 AACCTGTGAGAGCTATTTGAAATCATTTGCTTTTCAAGCTTACACAGGTTGGAGT 420  
 361 AACCTGTGAGAGCTATTTGAAATCATTTGCTTTTCAAGCTTACACAGGTTGGAGT 420  
 421 ATGCAAAAGCTATTAATTTTGGCAAAAAGGAAATTAACCTCTCTGACATCTTAAGAATG 480  
 421 ATGCAAAAGCTATTAATTTTGGCAAAAAGGAAATTAACCTCTCTGACATCTTAAGAATG 480  
 481 AAGTTTCTATCTCCAAAGTATGGGCTACAGAAACCGTCCAAAGAGCTTCTACAGAGTG 540  
 481 AAGTTTCTATCTCCAAAGTATGGGCTACAGAAACCGTCCAAAGAGCTTCTACAGAGTG 540  
 481 AAGTTTCTATCTCCAAAGTATGGGCTACAGAAACCGTCCAAAGAGCTTCTACAGAGTG 540  
 541 AACCCGAAATCTCTCTGAGAAACAGCTCAGTGTCCAGCTCTTAACCTTTGGA 600  
 541 AACCCGAAATCTCTCTGAGAAACAGCTCAGTGTCCAGCTCTTAACCTTTGGA 600  
 541 AACCCGAAATCTCTCTGAGAAACAGCTCAGTGTCCAGCTCTTAACCTTTGGA 600  
 601 CTGTGAGAACTCTGAGGACAAAGCAGCGATCAACCTCAAAAAGAGCTGTCTCATATTG 660  
 601 CTGTGAGAACTCTGAGGACAAAGCAGCGATCAACCTCAAAAAGAGCTGTCTCATATTG 660  
 661 AATTGGATCTATCTTCTGAGAGTACCGGTTAATAAGCAACTATTGCAAGTGGAG 720  
 661 AATTGGATCTATCTTCTGAGAGTACCGGTTAATAAGCAACTATTGCAAGTGGAG 720  
 721 ATCAGAATTTTGAATAATCACCCCTCAAGAAACAGGATGAATCAATTGGATTCTG 780  
 721 ATCAGAATTTTGAATAATCACCCCTCAAGAAACAGGATGAATCAATTGGATTCTG 780  
 781 CAAAAAAGGCTCTGTGGAATTTTCTGAGAGCGATGTAACTGCAATCAATCAATCAAC 840  
 781 CAAAAAAGGCTCTGTGGAATTTTCTGAGAGCGATGTAACTGCAATCAATCAATCAAC 840

841 CCAGTAATATGATTTGAACACCACTGAGAGCGGTGAGAGGATCCAGAAAGT 900  
 841 CCAGTAATATGATTTGAACACCACTGAGAGCGGTGAGAGGATCCAGAAAGT 900  
 901 ATCAGAGTAGTTCTGTTTCAAACTTGCATGTGGAGCCATGTGGCAAAATCTATGCGCA 960  
 901 ATCAGAGTAGTTCTGTTTCAAACTTGCATGTGGAGCCATGTGGCAAAATCTATGCGCA 960  
 901 ATCAGAGTAGTTCTGTTTCAAACTTGCATGTGGAGCCATGTGGCAAAATCTATGCGCA 960  
 961 GCTCATTCAGCATGAGAAACAGCAGTATTAATCACTCAAGAGCAAGATGATGTAGAA 1020  
 961 GCTCATTCAGCATGAGAAACAGCAGTATTAATCACTCAAGAGCAAGATGATGTAGAA 1020  
 1021 AGGCTGAATTCGTATTAAGCAACAGCCGCTGTGAGCAAGGAGCCATCAACAT 1080  
 1021 AGGCTGAATTCGTATTAAGCAACAGCCGCTGTGAGCAAGGAGCCATCAACAT 1080  
 1081 GGGCTGGAAGTAAAGAAACATGTATGTAGCGGACTCCAGCAGAAAAAAGGTAG 1140  
 1081 GGGCTGGAAGTAAAGAAACATGTATGTAGCGGACTCCAGCAGAAAAAAGGTAG 1140  
 1141 ATCTGAATCTGATCCCTGTGTGAGAGAAAGATGGAATAGCAGAAACTGCCATGCT 1200  
 1141 ATCTGAATCTGATCCCTGTGTGAGAGAAAGATGGAATAGCAGAAACTGCCATGCT 1200  
 1201 CAGAGAACTCTAGATATCTGAGAGTGTCTTGGATTAACACTAATAGCAGCATTCAGA 1260  
 1201 CAGAGAACTCTAGATATCTGAGAGTGTCTTGGATTAACACTAATAGCAGCATTCAGA 1260  
 1261 AAGTTAATGAGTGTGTTTCCAGAAATGATGACTGTAGTGTCTGATGACTCATGTATG 1320  
 1261 AAGTTAATGAGTGTGTTTCCAGAAATGATGACTGTAGTGTCTGATGACTCATGTATG 1320  
 1321 GGGAGTCTGAATCAAAATGCCAAAGTACTGATGTAATGAGCTTCTAAATGAGTATG 1380  
 1321 GGGAGTCTGAATCAAAATGCCAAAGTACTGATGTAATGAGCTTCTAAATGAGTATG 1380  
 1381 AATATTGTGTTCTTCCAGAAATGATGACTGTAGTGTCTGATGACTCATGTATG 1440  
 1381 AATATTGTGTTCTTCCAGAAATGATGACTGTAGTGTCTGATGACTCATGTATG 1440  
 1441 TATGTAAGTGAAGAGTTCATCCAAATGATGAGAGTAAATTTGAGCAAAATAT 1500  
 1441 TATGTAAGTGAAGAGTTCATCCAAATGATGAGAGTAAATTTGAGCAAAATAT 1500  
 1501 TTGGGAAACCTATGGAAGAGGCAAGCTTCCCACTTAAGCCATGTAAGTGAATC 1560  
 1501 TTGGGAAACCTATGGAAGAGGCAAGCTTCCCACTTAAGCCATGTAAGTGAATC 1560  
 1561 TTAATTATGAGACATTTGTTACTGAGCCACAGATATACAAAGGCTCCCTCACAAAT 1620  
 1561 TTAATTATGAGACATTTGTTACTGAGCCACAGATATACAAAGGCTCCCTCACAAAT 1620  
 1621 AATTAAAGGTAAGAGAGCTTACATGAGGCTTCACTGAGGATTTTATCAAGAAAG 1680  
 1621 AATTAAAGGTAAGAGAGCTTACATGAGGCTTCACTGAGGATTTTATCAAGAAAG 1680  
 1681 CAGATTTGGCAGTTCAAAAGACTCTGAAATGATTAACAGGAGCACTAACCAAGGAGC 1740  
 1681 CAGATTTGGCAGTTCAAAAGACTCTGAAATGATTAACAGGAGCACTAACCAAGGAGC 1740  
 1741 AGAATGGTCAAGTATGATATTAATTAATAGGTCATGAGAAATTAACCAAGGAGT 1800  
 1741 AGAATGGTCAAGTATGATATTAATTAATAGGTCATGAGAAATTAACCAAGGAGT 1800  
 1801 CTATTCAGATGAGAAATCTTAACCAATGATGATCACTGAGAAAGATGTGTTCA 1860  
 1801 CTATTCAGATGAGAAATCTTAACCAATGATGATCACTGAGAAAGATGTGTTCA 1860  
 1861 AAACGAAAGCTGAACTATAGCAGCAGTATAGCAATATGGAAGTCAATTAATATCC 1920  
 1861 AAACGAAAGCTGAACTATAGCAGCAGTATAGCAATATGGAAGTCAATTAATATCC 1920

QY	1921	ACAATTCAAAAGCACCCTAAAAAAGAT	TAGCGTGGAGAGAGAACTCTTCTACCGAGCATATTC	1980
Db	1921	ACAAATTCAAAAGCACCCTAAAAAGAT	TAGCGTGGAGAGAGAACTCTTCTACCGAGCATATTC	1980
QY	1981	ATGCGCTTGAACCTAGTACGAGTACAAAT	CTTAAGCGCCACTAATGTGCTCAATTTGCAAA	2040
Db	1981	ATGCGCTTGAACCTAGTACGAGTACAAAT	CTTAAGCGCCACTAATGTGCTCAATTTGCAAA	2040
QY	2041	TTGATAGTTGTTCTAGCAGTGAAGAT	TAAGAAAAAAGTTACAAACCAATATGCCAGTCA	2100
Db	2041	TTGATAGTTGTTCTAGCAGTGAAGAT	TAAGAAAAAAGTTACAAACCAATATGCCAGTCA	2100
QY	2101	GGCACAAGCAAAACCTACAACTCAT	TGGAAGGTAAAGAACTCTGCAACTGGAGCCAAAGA	2160
Db	2101	GGCACAAGCAAAACCTACAACTCAT	TGGAAGGTAAAGAACTCTGCAACTGGAGCCAAAGA	2160
QY	2161	GTAAACAAGCAAAATACAGACAAAT	TAAGAAAGCATGACAGTATCTTTCCAGAGCTGA	2220
Db	2161	GTAAACAAGCAAAATACAGACAAAT	TAAGAAAGCATGACAGTATCTTTCCAGAGCTGA	2220
QY	2221	AGTTAACAATATGCACCTGTTCTTTT	TACTAAGTGTTCAAATACAGTAACCTTAAGAAAT	2280
Db	2221	AGTTAACAATATGCACCTGTTCTTTT	TACTAAGTGTTCAAATACAGTAACCTTAAGAAAT	2280
QY	2281	TTGTCAATTCCTAGCCTCTCCAGAGA	AAAAAAGAAAGACAAACTAGAAACAGTTAAAGT	2340
Db	2281	TTGTCAATTCCTAGCCTCTCCAGAGA	AAAAAAGAAAGACAAACTAGAAACAGTTAAAGT	2340
QY	2341	CTAATAATATGCTCAACCCCAAGAT	CTCATGTTTAAGTGAAGAAAGGTTTGGCAAACTG	2400
Db	2341	CTAATAATATGCTCAACCCCAAGAT	CTCATGTTTAAGTGAAGAAAGGTTTGGCAAACTG	2400
QY	2401	AAAGATCTGTAGAGAGTAGCAAT	ATTTTCATTTGTAACCTGTAAGTATATGGAACACTGAG	2460
Db	2401	AAAGATCTGTAGAGAGTAGCAAT	ATTTTCATTTGTAACCTGTAAGTATATGGAACACTGAG	2460
QY	2461	AAAGTATCTCTTACTGTGAGAGTT	TAGCACTCTAAGGAGGCAAAAACAGAACCAATTAAT	2520
Db	2461	AAAGTATCTCTTACTGTGAGAGTT	TAGCACTCTAAGGAGGCAAAAACAGAACCAATTAAT	2520
QY	2521	GTGTAGTACAGTGTGCACGATTT	GGAAACCCCAAGGAGCAATTAATCATGGTTGTCCAAAG	2580
Db	2521	GTGTAGTACAGTGTGTGCACGATTT	GGAAACCCCAAGGAGCAATTAATCATGGTTGTCCAAAG	2580
QY	2581	ATAATAGAAATGACACAGAGGCT	TTAAGTATCCATTGGGACATGAAGTTAAACACAGTC	2640
Db	2581	ATAATAGAAATGACACAGAGGCT	TTAAGTATCCATTGGGACATGAAGTTAAACACAGTC	2640
QY	2641	GGGAACAAGCAATAGAAATGGAAG	CAAACTGAATGATCTCAGATTTATGGCAATACAT	2700
Db	2641	GGGAACAAGCAATAGAAATGGAAG	CAAACTGAATGATCTCAGATTTATGGCAATACAT	2700
QY	2701	TCAAGGTTTCAAGAGGCGCAGTAT	TGCTGCTGTTTTCAAAATCCAGGAATCCAGAAAGG	2760
Db	2701	TCAAGGTTTCAAGAGGCGCAGTAT	TGCTGCTGTTTTCAAAATCCAGGAATCCAGAAAGG	2760
QY	2761	AATGTGCAACAATCTCTGCCCCACT	CTGGGTCTTAAAGAAACAAAGTCCAAAGTCACTT	2820
Db	2761	AATGTGCAACAATCTCTGCCCCACT	CTGGGTCTTAAAGAAACAAAGTCCAAAGTCACTT	2820
QY	2821	TTGAATGTGAACAAAAGAGAAAT	TAAGAAATGAGTCTTAATATCAAGCTGTAC	2880
Db	2821	TTGAATGTGAACAAAAGAGAAAT	TAAGAAATGAGTCTTAATATCAAGCTGTAC	2880
QY	2881	AGACAGTAAATATACCTGACAGGCT	TCTGCTGGTTGTCANAAAGATTAAGCCATTGATA	2940
Db	2881	AGACAGTAAATATACCTGACAGGCT	TCTGCTGGTTGTCANAAAGATTAAGCCATTGATA	2940
QY	2941	ATGCCAAATGTAGTATCAAAAGAG	GCCTTAGGTTTGTCTATCTCATCTCAGTTCCAGAGGCA	3000
Db	2941	ATGCCAAATGTAGTATCAAAAGAG	GCCTTAGGTTTGTCTATCTCATCTCAGTTCCAGAGGCA	3000
QY	3001	ACGAAATCGACATTAATCTCCAAAT	AAACATGAGACTTTTACAAACCCATATATGCTATAC	3060

Db	3001	ACGAAACCTGGACTCATCTACTCCAAATAAACATGAGCTTTTACAAAACCCATATCTGTATAC	3066
OY	3061	CACCACTTTTCCCATCAAGTCATTTGTTAAACTAAATGTAGAAAAAATCTGCTAGAG	3120
Db	3061	CACCACTTTTCCCATCAAGTCATTTGTTAAACTAAATGTAGAAAAAATCTGCTAGAG	3120
OY	3121	AAACCTTTGAGAGACATTCATATGTCACCTGAAGAGAAATGGGAATAGAAATTCGCA	3180
Db	3121	AAACCTTTGAGAGACATTCATATGTCACCTGAAGAGAAATGGGAATAGAAATTCGCA	3180
OY	3181	GTACAGTGGCCAACTTAGCCCTAATTAACATTAGAAAAATGTTTTTAAAGAGCCACT	3240
Db	3181	GTACAGTGGCCAACTTAGCCCTAATTAACATTAGAAAAATGTTTTTAAAGAGCCACT	3240
OY	3241	CAGAAGATTTAATGAAGTAGTTCACAGTAAATGAAGTGGGCTCCAGTATTAATGAA	3300
Db	3241	CAGAAGATTTAATGAAGTAGTTCACAGTAAATGAAGTGGGCTCCAGTATTAATGAA	3300
OY	3301	TAGGTTCCAGTATGAAAAACATTCAGAAGAACTAGTAAGAAACAGAGGCCAAATTTGA	3360
Db	3301	TAGGTTCCAGTATGAAAAACATTCAGAAGAACTAGTAAGAAACAGAGGCCAAATTTGA	3360
OY	3361	ATGCTATGCTTAGATTAGGGGTTTTTGCAACCTGAGGTCTATTAACAAAGTCTTCTGGAA	3420
Db	3361	ATGCTATGCTTAGATTAGGGGTTTTTGCAACCTGAGGTCTATTAACAAAGTCTTCTGGAA	3420
OY	3421	GTAATTTGTAAGCATCTCTGAATAAAAAAGCAGAATATGAAGATAGTTCAGACTGTTA	3480
Db	3421	GTAATTTGTAAGCATCTCTGAATAAAAAAGCAGAATATGAAGATAGTTCAGACTGTTA	3480
OY	3481	ATACAGATTTCTCTCCATATCTGATTTAGATTAAGCAAGCCATATGGGAAGTAGTC	3540
Db	3481	ATACAGATTTCTCTCCATATCTGATTTAGATTAAGCAAGCCATATGGGAAGTAGTC	3540
OY	3541	ATGCATCTCAGGTTTGTCTGAGACACCTGATGACCTGTAGATATGCTGAATTAAGG	3600
Db	3541	ATGCATCTCAGGTTTGTCTGAGACACCTGATGACCTGTAGATATGCTGAATTAAGG	3600
OY	3601	AAAGTACTAGCTTTCTCTGAAAAATGACATTAAGGAAAGTTCGCTTTTATGCAAAACCG	3660
Db	3601	AAAGTACTAGCTTTCTCTGAAAAATGACATTAAGGAAAGTTCGCTTTTATGCAAAACCG	3660
OY	3661	TCCAGAAAGAGAGGAGCTTAGCAGAGATCTAGCCCTTACCCATACACATTTGGCTCAGG	3720
Db	3661	TCCAGAAAGAGAGGAGCTTAGCAGAGATCTAGCCCTTACCCATACACATTTGGCTCAGG	3720
OY	3721	GTTACCCGAAGAGGGGCCAAGAAATTAGAGTCTCGAAGAGAGACTTATCTAGTGAGATG	3780
Db	3721	GTTACCCGAAGAGGGGCCAAGAAATTAGAGTCTCGAAGAGAGACTTATCTAGTGAGATG	3780
OY	3781	AAGAGCTTCCCGTCTCCACACCTGTTATTGTTGAAGTAAACATATACCTTCTCAGT	3840
Db	3781	AAGAGCTTCCCGTCTCCACACCTGTTATTGTTGAAGTAAACATATACCTTCTCAGT	3840
OY	3841	CTACTAGGATAGACACCGTTGCTACCGAGTGTCTGCTAAGAACACAGAGAGAAATTTAT	3900
Db	3841	CTACTAGGATAGACACCGTTGCTACCGAGTGTCTGCTAAGAACACAGAGAGAAATTTAT	3900
OY	3901	TATCATTTGAAGATAGCTTAAATGACTCAGATTAACAGATTAATTTGCAAAAGCATCTC	3960
Db	3901	TATCATTTGAAGATAGCTTAAATGACTCAGATTAACAGATTAATTTGCAAAAGCATCTC	3960
OY	3961	AGGACATCACCTTATGAGGAGAAACAAATGTCTGAGCTGTGTTTCTTCAACAGTCA	4020
Db	3961	AGGACATCACCTTATGAGGAGAAACAAATGTCTGAGCTGTGTTTCTTCAACAGTCA	4020
OY	4021	GTGAATTTGAAGACTTGAAGTCAAAATTAACACACCCAGATCTTCTTGAATGGTCTT	4080
Db	4021	GTGAATTTGAAGACTTGAAGTCAAAATTAACACACCCAGATCTTCTTGAATGGTCTT	4080
OY	4081	CCAAACCAATGAGGCATAGTCTGAAGGCCAGGAGTTGTTGATGACAAAGAAATTTGG	4140
Db	4081	CCAAACCAATGAGGCATAGTCTGAAGGCCAGGAGTTGTTGATGACAAAGAAATTTGG	4140

Db 4081 CCAAAATGAGGATCAGTCTGAAAGCCAGGAGATTGCTGATGACAGAAATGG 4140  
 QY 4141 TTTCAGATGATGAGAAAGAGAGAGGCTTGGAAAGAAATATCAAGAGAGCAAGCA 4200  
 Db 4141 TTTCAGATGATGAGAAAGAGAGAGGCTTGGAAAGAAATATCAAGAGAGCAAGCA 4200  
 QY 4201 TGGATTTCAAACTTAGGTGAGAGCAGATCTGGGTGTGAGAGTAAACCAAGCTCTGAG 4260  
 Db 4201 TGGATTTCAAACTTAGGTGAGAGCAGATCTGGGTGTGAGAGTAAACCAAGCTCTGAG 4260  
 QY 4261 ACTGCTCAGGGCTATCTCTGAGAGTACATTTTAACCACTGACAGAGAGATACCATGC 4320  
 Db 4261 ACTGCTCAGGGCTATCTCTGAGAGTACATTTTAACCACTGACAGAGAGATACCATGC 4320  
 QY 4321 AACATTAACCTGATTAAGCTCCAGAGAAATGGCTGAAGTACAGCTGTGTAAGCAAC 4380  
 Db 4321 AACATTAACCTGATTAAGCTCCAGAGAAATGGCTGAAGTACAGCTGTGTAAGCAAC 4380  
 QY 4381 ATGGAGGAGGCTTCTTAACAGCTACCTTCCATCATTAAGTACTCTTGTGCTTGAAG 4440  
 Db 4381 ATGGAGGAGGCTTCTTAACAGCTACCTTCCATCATTAAGTACTCTTGTGCTTGAAG 4440  
 QY 4441 ACCGCGAATTCAGAGCAAGCAATCAAGAAAGCAAGTATTAACCTTCAAGAAAGTA 4500  
 Db 4441 ACCGCGAATTCAGAGCAAGCAATCAAGAAAGCAAGTATTAACCTTCAAGAAAGTA 4500  
 QY 4501 GTGAATTAACCTTAAAGCTCCAGAGAAATGGCTGAAGTACAGCTGTGTAAGCAAC 4560  
 Db 4501 GTGAATTAACCTTAAAGCTCCAGAGAAATGGCTGAAGTACAGCTGTGTAAGCAAC 4560  
 QY 4561 CAGATAGTTCTACAGTAAATTAAGAAACAGAGAGTGAAGAGTATCCCTCTTAAT 4620  
 Db 4561 CAGATAGTTCTACAGTAAATTAAGAAACAGAGAGTGAAGAGTATCCCTCTTAAT 4620  
 QY 4621 GCCCATCAATTAAGTATAGTGTGATGATGATGATGATGATGATGATGATGATGAT 4680  
 Db 4621 GCCCATCAATTAAGTATAGTGTGATGATGATGATGATGATGATGATGATGATGAT 4680  
 QY 4681 ACTACCATCTCAAGAGAGCTCATTAAAGTGTGATGATGATGATGATGATGATGATGAT 4740  
 Db 4681 ACTACCATCTCAAGAGAGCTCATTAAAGTGTGATGATGATGATGATGATGATGATGAT 4740  
 QY 4741 AGCTGGGCCACAGATTTGAGGAAACATCTTACTTCCCAAGGCAAGATCTAGAGGAA 4800  
 Db 4741 AGCTGGGCCACAGATTTGAGGAAACATCTTACTTCCCAAGGCAAGATCTAGAGGAA 4800  
 QY 4801 CCCCTTACCCTGAATCTGGAATCAGGCTCTCTGATGATGATGATGATGATGATGATGAT 4860  
 Db 4801 CCCCTTACCCTGAATCTGGAATCAGGCTCTCTGATGATGATGATGATGATGATGATGAT 4860  
 QY 4861 AAGACAGAGCCCGAGTCAAGTCTGTTGGCAATCACTTCAACCTTCAACCTTCAACCTT 4920  
 Db 4861 AAGACAGAGCCCGAGTCAAGTCTGTTGGCAATCACTTCAACCTTCAACCTTCAACCTT 4920  
 QY 4921 AAGTCCCAATTTGAAGTGTGAGATCTGCCAGAGTCCAGCTCTGCTCATCTACTG 4980  
 Db 4921 AAGTCCCAATTTGAAGTGTGAGATCTGCCAGAGTCCAGCTCTGCTCATCTACTG 4980  
 QY 4981 ATACTGCTGGGTATTAATGCAATGAAAGAGTGTGAGAGGAGAGAGGAGAGGAGAGGAG 5040  
 Db 4981 ATACTGCTGGGTATTAATGCAATGAAAGAGTGTGAGAGGAGAGAGGAGAGGAGAGGAG 5040  
 QY 5041 CTTCACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5100  
 Db 5041 CTTCACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5100  
 QY 5101 AATTTATCTGCTGTACAGATTTGCCAGAGAAACACACATCACTTAACTAATTAATTA 5160  
 Db 5101 AATTTATCTGCTGTACAGATTTGCCAGAGAAACACACATCACTTAACTAATTAATTA 5160  
 QY 5161 CTGAAGAGACTACTGCTGTGATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5220  
 Db 5161 CTGAAGAGACTACTGCTGTGATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5220

QY 5221 TGAATATTTTCAGGAATTTGGGGAGAGAAATGGGTAGTATGCTATTTCTGGTGACCC 5280  
 Db 5221 TGAATATTTTCAGGAATTTGGGGAGAGAAATGGGTAGTATGCTATTTCTGGTGACCC 5280  
 QY 5281 ACTCTATTAAAGAGAGAAATGCTGAATGAGCATGATTTTGAAGTCAAGAGAGATGTCG 5340  
 Db 5281 ACTCTATTAAAGAGAGAAATGCTGAATGAGCATGATTTTGAAGTCAAGAGAGATGTCG 5340  
 QY 5341 TCAATGAGAGAAACCAAGAGTCCAAAGCGAGAGAGAGATGCCAGAGAGAGATCT 5400  
 Db 5341 TCAATGAGAGAAACCAAGAGTCCAAAGCGAGAGAGAGATGCCAGAGAGAGATCT 5400  
 QY 5401 TCAGGGGGCTAGAAATCTGCTTATGAGGCGCTTCAACCAATGCCAGATCAACTTCG 5460  
 Db 5401 TCAGGGGGCTAGAAATCTGCTTATGAGGCGCTTCAACCAATGCCAGATCAACTTCG 5460  
 QY 5461 AATGATGATGAGAGTGTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5520  
 Db 5461 AATGATGATGAGAGTGTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5520  
 QY 5521 GCACAGGTGTCCACCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5580  
 Db 5521 GCACAGGTGTCCACCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5580  
 QY 5581 TCCATGCAATTTGGGAGATGCTGAGAGGACCTGTGCTGCTGCTGCTGCTGCTGCTGCT 5640  
 Db 5581 TCCATGCAATTTGGGAGATGCTGAGAGGACCTGTGCTGCTGCTGCTGCTGCTGCTGCT 5640  
 QY 5641 GTGTAGCACTTACAGTGTCCAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 5700  
 Db 5641 GTGTAGCACTTACAGTGTCCAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 5700  
 QY 5701 GCCACTACTGA 5711  
 Db 5701 GCCACTACTGA 5711

RESULT 12  
 LOCUS 176943 5914 bp DNA linear PAT 03-APR-1998  
 DEFINITION Sequence 1 from patent US 5693473.  
 ACCESSION 176943  
 VERSION 176943.1 GI:3013097  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 5914)  
 AUTHORS Nakamura, Y., Shattuck-Eidens, D. M., Simard, J., Durocher, F., Emi, M. and  
 TITLE Linked breast and ovarian cancer susceptibility gene  
 JOURNAL Patent: US 5693473-A 1 02-DEC-1997;  
 FEATURES  
 source Location/Qualifiers  
 1..5914  
 /organism="unknown"  
 BASE COUNT 2006 a 1156 c 1316 g 1436 t  
 ORIGIN  
 Query Match 99.9%; Score 5707.8; DB 6; Length 5914;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 5709; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGCTGCTGAGACTTCTGAGAGCCCGGACAGAGCTGTGGGTTTCTGAGATTAAGTGGCC 60  
 Db 1 AGCTGCTGAGACTTCTGAGAGCCCGGACAGAGCTGTGGGTTTCTGAGATTAAGTGGCC 60  
 QY 61 CCTGGCTCAGAGAGGCTTCAACCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
 Db 61 CCTGGCTCAGAGAGGCTTCAACCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
 QY 121 TGGATTTATCTGCTCTTGGGCTTGAAGAGTAAAGATGCTAATTAATGCTATGAGAGAA 180  
 Db 121 TGGATTTATCTGCTCTTGGGCTTGAAGAGTAAAGATGCTAATTAATGCTATGAGAGAA 180

Db 121 TGCATTTATCTGCTCTTCCGCTTGAGAGATACAAAATGTCATTAAATGCTATGAGAAAA 180  
QY 181 TCTTAGAGTGTCCATCTGTCTGGAGTGTATCAAGAAACCTGTCTCCAAAGTGTGACC 240  
Db 181 TCTTAGAGTGTCCATCTGTCTGGAGTGTATCAAGAAACCTGTCTCCAAAGTGTGACC 240  
QY 241 ACATATTTTGAATTTTGGCATGCTGAAACCTTCTACACAGACAGAAAGGCCCTTACAGT 300  
Db 241 ACATATTTTGAATTTTGGCATGCTGAAACCTTCTACACAGACAGAAAGGCCCTTACAGT 300  
QY 301 GTCCCTTATGTAAGATGATATTAACCAAAAGAGCCCTCAAGAAAGTACGAGATTATATC 360  
Db 301 GTCCCTTATGTAAGATGATATTAACCAAAAGAGCCCTCAAGAAAGTACGAGATTATATC 360  
QY 361 AACTTGTGAAGACCTATTTGAATAATCATTTTGTGCTTTTACCTTGACACAGTTTGGAGT 420  
Db 361 AACTTGTGAAGACCTATTTGAATAATCATTTTGTGCTTTTACCTTGACACAGTTTGGAGT 420  
QY 421 ATGCAAAACAGCTATTAATTTTGCAGAAAATTAACCTTCTGACATCTTAAAGATG 480  
Db 421 ATGCAAAACAGCTATTAATTTTGCAGAAAATTAACCTTCTGACATCTTAAAGATG 480  
QY 481 AAGTTTATCATCCAAAGTATGGGCTACAGAAACCCGTCAGAAAAGTCTTACAGAGT 540  
Db 481 AAGTTTATCATCCAAAGTATGGGCTACAGAAACCCGTCAGAAAAGTCTTACAGAGT 540  
QY 541 AACCCGAAATTCCTTCTTGAGAGAAACAGCTCAGTGTCCAACTCTTAACCTTGGAA 600  
Db 541 AACCCGAAATTCCTTCTTGAGAGAAACAGCTCAGTGTCCAACTCTTAACCTTGGAA 600  
QY 601 CTGTGAGACTCTGAGAGCAAAAGCGGATACAACTCAAAAGAGCTGTCTACATTTG 660  
Db 601 CTGTGAGACTCTGAGAGCAAAAGCGGATACAACTCAAAAGAGCTGTCTACATTTG 660  
QY 661 AATTGGATCGATCTTCTGAAAGATACCGTTAATAGGCAACTATTTGCGAGTGGAG 720  
Db 661 AATTGGATCGATCTTCTGAAAGATACCGTTAATAGGCAACTATTTGCGAGTGGAG 720  
QY 721 ATCAAGAATTTGTACAATCAACCCCTCAAGAACAGGATGAATCAGTTTGGATTCTG 780  
Db 721 ATCAAGAATTTGTACAATCAACCCCTCAAGAACAGGATGAATCAGTTTGGATTCTG 780  
QY 781 CAAAAAAGGCTGCTGTGGAATTTTCTGAGACGAGTGAACAAATACGAAATCATATC 840  
Db 781 CAAAAAAGGCTGCTGTGGAATTTTCTGAGACGAGTGAACAAATACGAAATCATATC 840  
QY 841 CCAGTAAATATGATTTGAACACCACTGAGAGCGTGCAGCTGAGAGGCATCCAGAAAAT 900  
Db 841 CCAGTAAATATGATTTGAACACCACTGAGAGCGTGCAGCTGAGAGGCATCCAGAAAAT 900  
QY 901 ATCAGGATAGTCTGTTTCAAACTTGCATGTGAGCCATGTGCACAAAATACTCATGCA 960  
Db 901 ATCAGGATAGTCTGTTTCAAACTTGCATGTGAGCCATGTGCACAAAATACTCATGCA 960  
QY 961 GCTCATTACGATGAGAACAGCAGTTTATTCTACATAAGACAGATGATAGAAA 1020  
Db 961 GCTCATTACGATGAGAACAGCAGTTTATTCTACATAAGACAGATGATAGAAA 1020  
QY 1021 AGGCTGAATCTGTAATTAAGCAAAACAGCCTGGCTTAGAAGAGGCAACATTAACAGAT 1080  
Db 1021 AGGCTGAATCTGTAATTAAGCAAAACAGCCTGGCTTAGAAGAGGCAACATTAACAGAT 1080  
QY 1081 GGGCTGGAAGTAAAGAAACATGTAATGATAGGGGAGCTCCAGACAGAAAAAGGTAG 1140  
Db 1081 GGGCTGGAAGTAAAGAAACATGTAATGATAGGGGAGCTCCAGACAGAAAAAGGTAG 1140  
QY 1141 ATCTGAATGCTGATCCCTGTGTGAGAGAAAGATGGAATTAAGCAAAATGCAATGCT 1200  
Db 1141 ATCTGAATGCTGATCCCTGTGTGAGAGAAAGATGGAATTAAGCAAAATGCAATGCT 1200  
QY 1201 CAGAGATCTTACAGATGCTGAGAGATGCTCTTGGATTAACACTTAATAGCAGATTGAGA 1260  
Db 1201 CAGAGATCTTACAGATGCTGAGAGATGCTCTTGGATTAACACTTAATAGCAGATTGAGA 1260

QY 1261 AAGTTAATGATGTTTTCCAGAAAGTATGAACTGTTAGTTTGTATGACTCACANGATG 1320  
Db 1261 AAGTTAATGATGTTTTCCAGAAAGTATGAACTGTTAGTTTGTATGACTCACANGATG 1320  
QY 1321 GGGAGTCTGAATCAAAATCCCAAGTACGTGATGTTTGGACGTTCTTAATAGGTAGATG 1380  
Db 1321 GGGAGTCTGAATCAAAATCCCAAGTACGTGATGTTTGGACGTTCTTAATAGGTAGATG 1380  
QY 1381 AATATTCGTGTTCTTCAGAGAAATATGACTTACGGCAGAGATCCCTCATGAGCCTTTAA 1440  
Db 1381 AATATTCGTGTTCTTCAGAGAAATATGACTTACGGCAGAGATCCCTCATGAGCCTTTAA 1440  
QY 1441 TATGTAAGATGAAAGAGTTTCTACTCCAATCAGTAGAGATTAATTTGAAGACAAAATAT 1500  
Db 1441 TATGTAAGATGAAAGAGTTTCTACTCCAATCAGTAGAGATTAATTTGAAGACAAAATAT 1500  
QY 1501 TTGGGAAAACCTATCGAGAGAGCAGCAGCCCTCCCACTTAAGCATGTAACTGAAATTC 1560  
Db 1501 TTGGGAAAACCTATCGAGAGAGCAGCAGCCTCCCACTTAAGCATGTAACTGAAATTC 1560  
QY 1561 TAATTATAGAGCATTTGTTACTGAGCCACAGATTAATACAAAGCCCTCCACAAATA 1620  
Db 1561 TAATTATAGAGCATTTGTTACTGAGCCACAGATTAATACAAAGCCCTCCACAAATA 1620  
QY 1621 AATTAAAGCGTAAGAGAGACCTACATCAGGCTTCTATCCTGAGGATTTTATCAAGAAAG 1680  
Db 1621 AATTAAAGCGTAAGAGAGACCTACATCAGGCTTCTATCCTGAGGATTTTATCAAGAAAG 1680  
QY 1681 CAGATTTTGGCATTCAAAAGACCTCTGAATAATATCAAGGAACTTAACCAACGAGAC 1740  
Db 1681 CAGATTTTGGCATTCAAAAGACCTCTGAATAATATCAAGGAACTTAACCAACGAGAC 1740  
QY 1741 AGAATGGCAAGCATGAATTAATTAATGCTGCTATGAGATTAACAAAGAGTAT 1800  
Db 1741 AGAATGGCAAGCATGAATTAATTAATGCTGCTATGAGATTAACAAAGAGTAT 1800  
QY 1801 CTATTGAGAAATGAGAAAATCTTAACCAATAGAACTACGCAAAAAGAAATCTGCTTTCA 1860  
Db 1801 CTATTGAGAAATGAGAAAATCTTAACCAATAGAACTACGCAAAAAGAAATCTGCTTTCA 1860  
QY 1861 AAACGAAAGCTGAACCTTAAGACGAGTAAAGCAATATGAACTGCAATTAATATATCC 1920  
Db 1861 AAACGAAAGCTGAACCTTAAGACGAGTAAAGCAATATGAACTGCAATTAATATATCC 1920  
QY 1921 ACAATTCAAAAGCCTTAAGAAAGATGAGTGAAGGAGAGTCTTACACAGGCAATTC 1980  
Db 1921 ACAATTCAAAAGCCTTAAGAAAGATGAGTGAAGGAGAGTCTTACACAGGCAATTC 1980  
QY 1981 ATGCGCTTGAACATGATAGTAAATCTAAGCCACCTAATTTGATGAAATTCGAAA 2040  
Db 1981 ATGCGCTTGAACATGATAGTAAATCTAAGCCACCTAATTTGATGAAATTCGAAA 2040  
QY 2041 TTGATAGTGTGTTAGCAGTGAAGAGATTAAGAAAAAGTCAACCAATGCCAGTCA 2100  
Db 2041 TTGATAGTGTGTTAGCAGTGAAGAGATTAAGAAAAAGTCAACCAATGCCAGTCA 2100  
QY 2101 GGCACAGCAGAAACCTTCAACCTCATGGAAGTAAAGACCTGCACTGAGGCAAGAAAGA 2160  
Db 2101 GGCACAGCAGAAACCTTCAACCTCATGGAAGTAAAGACCTGCACTGAGGCAAGAAAGA 2160  
QY 2161 GTAACAAGCCCAATGAACAGACAGTAAAGAGACATGACAGTATCTTCCAGAGCTGA 2220  
Db 2161 GTAACAAGCCCAATGAACAGACAGTAAAGAGACATGACAGTATCTTCCAGAGCTGA 2220  
QY 2221 AGTTAACAATGCACTGGTCTTTTACTAGTGTTCCTCAATACACAGTAAAGAT 2280  
Db 2221 AGTTAACAATGCACTGGTCTTTTACTAGTGTTCCTCAATACACAGTAAAGAT 2280  
QY 2281 TTGTCAATCTTACGCTTCCAGAGAGAGAAAAAGCAAGCAAGTAAAGTGT 2340  
Db 2281 TTGTCAATCTTACGCTTCCAGAGAGAGAAAAAGCAAGCAAGTAAAGTGT 2340



QY 2341 CTAATATGCTGAGAACCCCAAGATCTCATGTTAAGTGAGAGAGGGTTTTGCAACGTG 2400  
 |||||  
 Db 2341 CTAATATGCTGAGAACCCCAAGATCTCATGTTAAGTGAGAGAGGGTTTTGCAACGTG 2400  
 QY 2401 AAAGATCTGTAGAGAGTAGACATATTTCATTGTGTACCTGTACTATTATATGCACTCAGG 2460  
 |||||  
 Db 2401 AAAGATCTGTAGAGAGTAGACATATTTCATTGTGTACCTGTACTATTATATGCACTCAGG 2460  
 QY 2461 AAAGATCTGTTACTGTAGAGAGTAGACATCTTAGGAGGCAAAAACAGAACCAATTAAT 2520  
 |||||  
 Db 2461 AAAGATCTGTTACTGTAGAGAGTAGACATCTTAGGAGGCAAAAACAGAACCAATTAAT 2520  
 QY 2521 GTGTAGTCAAGTGTGACCATTTTGAAGACCCCAAGGAGTAATTCATGTTGTTCCAAAG 2580  
 |||||  
 Db 2521 GTGTAGTCAAGTGTGACCATTTTGAAGACCCCAAGGAGTAATTCATGTTGTTCCAAAG 2580  
 QY 2581 ATAATAGAAATGACACAGAGGCTTTAAGTATCCATTGGGACATAGTTAACCCAGTC 2640  
 |||||  
 Db 2581 ATAATAGAAATGACACAGAGGCTTTAAGTATCCATTGGGACATAGTTAACCCAGTC 2640  
 QY 2641 GGAATACAGCATTAAGAAATGAGAAAGTAAGTATGATGCTGATGATTGAGATACAT 2700  
 |||||  
 Db 2641 GGAATACAGCATTAAGAAATGAGAAAGTAAGTATGATGCTGATGATTGAGATACAT 2700  
 QY 2701 TCAGGTTTCAAGGCGCAGTCATTGCTGTGTTTCAATCCAGAAATGACAGAGAG 2760  
 |||||  
 Db 2701 TCAGGTTTCAAGGCGCAGTCATTGCTGTGTTTCAATCCAGAAATGACAGAGAG 2760  
 QY 2761 AATGTGCAACATTTCTGCTGCTGCTGCTGCTTAAAGAAACAAAGTCCAAAGTCATT 2820  
 |||||  
 Db 2761 AATGTGCAACATTTCTGCTGCTGCTGCTGCTTAAAGAAACAAAGTCCAAAGTCATT 2820  
 QY 2821 TTGAATGTGACAAAGAAAGAAATCAAGAAAGATGATGATATATCAAGCTGTATC 2880  
 |||||  
 Db 2821 TTGAATGTGACAAAGAAAGAAATCAAGAAAGATGATGATATATCAAGCTGTATC 2880  
 QY 2881 AGACAGTTAATATATCACTGAGGCTTCTGTGTTGTCAGAAAGATAGCCAGTTGATA 2940  
 |||||  
 Db 2881 AGACAGTTAATATATCACTGAGGCTTCTGTGTTGTCAGAAAGATAGCCAGTTGATA 2940  
 QY 2941 ATGCCAATGTATATCAAGAGAGGCTTAGGTTTGTCTATCATCTCAGTTCAAGAGCA 3000  
 |||||  
 Db 2941 ATGCCAATGTATATCAAGAGAGGCTTAGGTTTGTCTATCATCTCAGTTCAAGAGCA 3000  
 QY 3001 AGCAATGAGACTCATTAATTAACATTAACATTAACATTAACATTAACATTAACATTAAC 3060  
 |||||  
 Db 3001 AGCAATGAGACTCATTAATTAACATTAACATTAACATTAACATTAACATTAACATTAAC 3060  
 QY 3061 CACCACTTTTCCATCAAGTCATTGTTAAACTAAATGTAGAAATATGCTAGAGG 3120  
 |||||  
 Db 3061 CACCACTTTTCCATCAAGTCATTGTTAAACTAAATGTAGAAATATGCTAGAGG 3120  
 QY 3121 AAAATTTTGTAGAGACATTCATGTACCTGAAAGGAAATGGGAAATGAGAAATTCCTCA 3180  
 |||||  
 Db 3121 AAAATTTTGTAGAGACATTCATGTACCTGAAAGGAAATGGGAAATGAGAAATTCCTCA 3180  
 QY 3181 GTACAGTACGACATTAAGGCTTAATTAACATTAAGAGAAATGTTTAAAGAGGACAGT 3240  
 |||||  
 Db 3181 GTACAGTACGACATTAAGGCTTAATTAACATTAAGAGAAATGTTTAAAGAGGACAGT 3240  
 QY 3241 CAAGCAATTAATTAAGTAGGTTCCAGTACTAATGAAGTGGGCTCCAGTATTAATGA 3300  
 |||||  
 Db 3241 CAAGCAATTAATTAAGTAGGTTCCAGTACTAATGAAGTGGGCTCCAGTATTAATGA 3300  
 QY 3301 TAGGTTTCCAGTATTAAGTAACATTCAGAGAACTAGTGAAGAAACAGAGGCTCAAAATGA 3360  
 |||||  
 Db 3301 TAGGTTTCCAGTATTAAGTAACATTCAGAGAACTAGTGAAGAAACAGAGGCTCAAAATGA 3360  
 QY 3361 ATGCTATGCTTAGATTAGGGGTTTGGCAACCTGAGGCTATATTAAGAAAGCTTCTCTGGA 3420  
 |||||  
 Db 3361 ATGCTATGCTTAGATTAGGGGTTTGGCAACCTGAGGCTATATTAAGAAAGCTTCTCTGGA 3420  
 QY 3421 GTAATTTGATGACATCTGAAATTAAGAAAGCAAGATATGAGAAAGTACGACTGTTA 3480

Db 3421 GTAATTTGATGACATCTGAAATTAAGAAAGCAAGATATGAAAGTAGTTACAGCTGTA 3480  
 |||||  
 QY 3481 ATACAGATTTCTCTCCATATCTGATTTCAGATTAACCTTAAAGAGCTATAGGAGTAGTC 3540  
 |||||  
 Db 3481 ATACAGATTTCTCTCCATATCTGATTTCAGATTAACCTTAAAGAGCTATAGGAGTAGTC 3540  
 QY 3541 ATGCAATCAGGTTTGTCTGTGACACCGTATGACCTGTATGATGATGATGATGATGATGATG 3600  
 |||||  
 Db 3541 ATGCAATCAGGTTTGTCTGTGACACCGTATGACCTGTATGATGATGATGATGATGATGATG 3600  
 QY 3601 AAGATACATGTTTCTGTAATAATGACATTAAGAAAGTCTGCTGTTTAAAGCAAGCG 3660  
 |||||  
 Db 3601 AAGATACATGTTTCTGTAATAATGACATTAAGAAAGTCTGCTGTTTAAAGCAAGCG 3660  
 QY 3661 TCCAGAAAGGAGGCTTAGCAGAGAGTCCAGCCCTTACCCATACACATTTGGCTCAGG 3720  
 |||||  
 Db 3661 TCCAGAAAGGAGGCTTAGCAGAGAGTCCAGCCCTTACCCATACACATTTGGCTCAGG 3720  
 QY 3721 GTTACCGAAGAGGCGCCAAAGAAATTAAGTCTCAGAGAGAACTATCTAGTAGAGATG 3780  
 |||||  
 Db 3721 GTTACCGAAGAGGCGCCAAAGAAATTAAGTCTCAGAGAGAACTATCTAGTAGAGATG 3780  
 QY 3781 AAGAGCTTCCCTGCTTCCAAACACTTGTATTGTAAGTAAACAATATACCTTCTCAGT 3840  
 |||||  
 Db 3781 AAGAGCTTCCCTGCTTCCAAACACTTGTATTGTAAGTAAACAATATACCTTCTCAGT 3840  
 QY 3841 CTACTAGGCTATGACACCGCTGCTACCAGTGTCTGTCTAAGAACACAGAGGAAATTTAT 3900  
 |||||  
 Db 3841 CTACTAGGCTATGACACCGCTGCTACCAGTGTCTGTCTAAGAACACAGAGGAAATTTAT 3900  
 QY 3901 TATCATGAGAAATAGCTTAATTAATGATGATGATGATGATGATGATGATGATGATGATG 3960  
 |||||  
 Db 3901 TATCATGAGAAATAGCTTAATTAATGATGATGATGATGATGATGATGATGATGATGATG 3960  
 QY 3961 AGGAATCATCACCTTAGTGAAGAAACAAATGTTGCTAGCTGTTGTTCTTCCACAGTGA 4020  
 |||||  
 Db 3961 AGGAATCATCACCTTAGTGAAGAAACAAATGTTGCTAGCTGTTGTTCTTCCACAGTGA 4020  
 QY 4021 GTGAATTTGAGAACTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 4080  
 |||||  
 Db 4021 GTGAATTTGAGAACTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 4080  
 QY 4081 CCAACCAATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4140  
 |||||  
 Db 4081 CCAACCAATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4140  
 QY 4141 TTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4200  
 |||||  
 Db 4141 TTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4200  
 QY 4201 TGGATTTCAACTTAGTGAAGAGCATCTGGGTTGAGAGTGAAGAAACAGGCTCTGTAAG 4260  
 |||||  
 Db 4201 TGGATTTCAACTTAGTGAAGAGCATCTGGGTTGAGAGTGAAGAAACAGGCTCTGTAAG 4260  
 QY 4261 ACTGCTCAGGCTATCTCTCTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 4320  
 |||||  
 Db 4261 ACTGCTCAGGCTATCTCTCTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 4320  
 QY 4321 AACATTAACCTGATTAAGCTCCAGAGGAAATGGCTGAAGTGAAGTGAAGTGAAGTGAAG 4380  
 |||||  
 Db 4321 AACATTAACCTGATTAAGCTCCAGAGGAAATGGCTGAAGTGAAGTGAAGTGAAGTGAAG 4380  
 QY 4381 ATGGAAGCAGGCTTCTTAACAGCTACCTTCCATATTAAGTGAAGTGAAGTGAAGTGAAG 4440  
 |||||  
 Db 4381 ATGGAAGCAGGCTTCTTAACAGCTACCTTCCATATTAAGTGAAGTGAAGTGAAGTGAAG 4440  
 QY 4441 AECTGCAATTCACAAACAAAGCAGATCAGAAAGAGCTATTAACCTTACAGAAAGATA 4500  
 |||||  
 Db 4441 AECTGCAATTCACAAACAAAGCAGATCAGAAAGAGCTATTAACCTTACAGAAAGATA 4500  
 QY 4441 AECTGCAATTCACAAACAAAGCAGATCAGAAAGAGCTATTAACCTTACAGAAAGATA 4500  
 |||||  
 Db 4441 AECTGCAATTCACAAACAAAGCAGATCAGAAAGAGCTATTAACCTTACAGAAAGATA 4500  
 QY 4501 GTGAATTAACCTTATAAGCCAGAAATCCAGAGGCTTCTGCTGACAAAGTGAAGTGTG 4560  
 |||||





Db 541 AACCCGAAATCTCTCCCTGAGGAAACAGTCTCAGTGTCCACTCTCTTAACCTTGAA 600  
OY 601 CTGTGAGAACTCTGAGACAAAGACGGATACACACTCAAAAGACGTCTGTACATG 660  
Db 601 CTGTGAGAACTCTGAGACAAAGACGGATACACACTCAAAAGACGTCTGTACATG 660  
OY 661 AATTGGAGTCTGATCTCTCTGAGATACCGTTAATAGGCAACTTATTGCACTGTGGAG 720  
Db 661 AATTGGAGTCTGATCTCTCTGAGATACCGTTAATAGGCAACTTATTGCACTGTGGAG 720  
OY 721 ATCAGAAATGTACAAATACCCCTCAAGAACCGAGGATGAATTCAGTTGATTCG 780  
Db 721 ATCAGAAATGTACAAATACCCCTCAAGAACCGAGGATGAATTCAGTTGATTCG 780  
OY 781 CAAAAAGAGCTCTGTGATTTCTGAGACGATGTAAACAATCTGACATCATCAAC 840  
Db 781 CAAAAAGAGCTCTGTGATTTCTGAGACGATGTAAACAATCTGACATCATCAAC 840  
OY 841 CCAGTAATATGATTTGAAACACCACTGAGAACGCTGAGAGGATCCAGAAAGT 900  
Db 841 CCAGTAATATGATTTGAAACACCACTGAGAACGCTGAGAGGATCCAGAAAGT 900  
OY 901 ATCAGAGTATCTCTCTCAAACTTCATGTGAGACCATGTGACAAATCTCATGCCA 960  
Db 901 ATCAGAGTATCTCTCTCAAACTTCATGTGAGACCATGTGACAAATCTCATGCCA 960  
OY 961 GCTCATTCACGATGAGAACACGAGTTTATTACTACTAATACAGAAATGATGAGAA 1020  
Db 961 GCTCATTCACGATGAGAACACGAGTTTATTACTACTAATACAGAAATGATGAGAA 1020  
OY 1021 AGGCTGAATCTGTATAAAGCAAAAGCCCTGGCTTAGCAGAGGACCAATTAACAGT 1080  
Db 1021 AGGCTGAATCTGTATAAAGCAAAAGCCCTGGCTTAGCAGAGGACCAATTAACAGT 1080  
OY 1081 GGGCTGGAAGTAAAGAAACATGTATGATGAGCGGACCTCCAGCAGAAAAAGGTAG 1140  
Db 1081 GGGCTGGAAGTAAAGAAACATGTATGATGAGCGGACCTCCAGCAGAAAAAGGTAG 1140  
OY 1141 ATCTGAATGCTGATCCCTGTGTGAGAAAAAGATGAATAGAGAAACCTGCATGCT 1200  
Db 1141 ATCTGAATGCTGATCCCTGTGTGAGAAAAAGATGAATAGAGAAACCTGCATGCT 1200  
OY 1201 CAGAGATCTAGAGATCTGAGAGATGTTCTTGGATTAACACTTAATAGCAGCATTCAG 1260  
Db 1201 CAGAGATCTAGAGATCTGAGAGATGTTCTTGGATTAACACTTAATAGCAGCATTCAG 1260  
OY 1261 AAGTTAATGAGTGGTTTCCAGAAAGTGAACCTGTTAGGTTCTGATGACATGATG 1320  
Db 1261 AAGTTAATGAGTGGTTTCCAGAAAGTGAACCTGTTAGGTTCTGATGACATGATG 1320  
OY 1321 GGGAGTCTGAATCAATGCGAAAGTGAATGATGATGAGGTTCTTAATGAGTAAAG 1380  
Db 1321 GGGAGTCTGAATCAATGCGAAAGTGAATGATGATGAGGTTCTTAATGAGTAAAG 1380  
OY 1381 AATATTTCTGTTCTTCAGAGAAATAGACTTACTGCCAGTGATCCTCATGAGGCTTAA 1440  
Db 1381 AATATTTCTGTTCTTCAGAGAAATAGACTTACTGCCAGTGATCCTCATGAGGCTTAA 1440  
OY 1441 TATGTAAAAGTGAAGAGTTCACCTCAATCAGTAGAGATTAATTTGAAGCAAAATAT 1500  
Db 1441 TATGTAAAAGTGAAGAGTTCACCTCAATCAGTAGAGATTAATTTGAAGCAAAATAT 1500  
OY 1501 TTGGGAAAAACCTATGGAAGAGGCAAGCTCCCACTTAAGCCATGATGAATGAAATC 1560  
Db 1501 TTGGGAAAAACCTATGGAAGAGGCAAGCTCCCACTTAAGCCATGATGAATGAAATC 1560  
OY 1561 TAATTTATGAGAGCAATTTGTACTGAGCCACAGATTAACAGAGGCTCCCTCACAAATA 1620  
Db 1561 TAATTTATGAGAGCAATTTGTACTGAGCCACAGATTAACAGAGGCTCCCTCACAAATA 1620  
OY 1621 AATTAAGCGTAAAGAGAGCTACATGAGGCTTCATCTGAGAGATTTTATCAAGAAAG 1680  
Db 1621 AATTAAGCGTAAAGAGAGCTACATGAGGCTTCATCTGAGAGATTTTATCAAGAAAG 1680

OY 1681 CAGATTTGGCAGTTCAAAAAGACTCTCTGAATGATAAATCAGGAACTTAACCAAGGAGC 1740  
Db 1681 CAGATTTGGCAGTTCAAAAAGACTCTCTGAATGATAAATCAGGAACTTAACCAAGGAGC 1740  
OY 1741 AGAATGTCAGTATGTAATTTACTAATAGTGTATGAGATTAATAACAAAGGTGAT 1800  
Db 1741 AGAATGTCAGTATGTAATTTACTAATAGTGTATGAGATTAATAACAAAGGTGAT 1800  
OY 1801 CTATTCAGATGAGAAAAATCTAACCCTAAGTAACTGACAAAAAGATCTGCTTCA 1860  
Db 1801 CTATTCAGATGAGAAAAATCTAACCCTAAGTAACTGACAAAAAGATCTGCTTCA 1860  
OY 1861 AAACGAAAGCTGAACTTATTAAGCAGCATATAGCAATTAAGCACTGAATTAATATC 1920  
Db 1861 AAACGAAAGCTGAACTTATTAAGCAGCATATAGCAATTAAGCACTGAATTAATATC 1920  
OY 1921 ACAATTCAAAAGCACCTTAANAAGATAGGCTGAGAGAGAGTCTTCTACAGGCATATTC 1980  
Db 1921 ACAATTCAAAAGCACCTTAANAAGATAGGCTGAGAGAGAGTCTTCTACAGGCATATTC 1980  
OY 1981 ATGCGCTGAACTAGTACGATAGAAATCTAAGCCCACTAATTTGTAATGCAAA 2040  
Db 1981 ATGCGCTGAACTAGTACGATAGAAATCTAAGCCCACTAATTTGTAATGCAAA 2040  
OY 2041 TTGATAGTTGTTCTAGCAGTGAAGATTAAGAAAAAAGTAAACCAATGCGAGTCA 2100  
Db 2041 TTGATAGTTGTTCTAGCAGTGAAGATTAAGAAAAAAGTAAACCAATGCGAGTCA 2100  
OY 2101 GGCACAGCAGAAACCTACAACTCATGGAAGTAAAGAACCTCACTGAGCCAGAAAGA 2160  
Db 2101 GGCACAGCAGAAACCTACAACTCATGGAAGTAAAGAACCTCACTGAGCCAGAAAGA 2160  
OY 2161 GTAAACGCGAAATGAACAGACAACTAAAGACATGACAGTATCTTCCAGAGCTGA 2220  
Db 2161 GTAAACGCGAAATGAACAGACAACTAAAGACATGACAGTATCTTCCAGAGCTGA 2220  
OY 2221 AGTTAAACAAATGACCTGCTTCTTACTAATGTTCAAAATACAGTAAAGAAAT 2280  
Db 2221 AGTTAAACAAATGACCTGCTTCTTACTAATGTTCAAAATACAGTAAAGAAAT 2280  
OY 2281 TTGTCAATCTAGCCTTCCAGAGAGAAAGAAAGAACTAGAAACAGTTAAAGTGT 2340  
Db 2281 TTGTCAATCTAGCCTTCCAGAGAGAAAGAAAGAAAGAACTAGAAACAGTTAAAGTGT 2340  
OY 2341 CTAAATTAATGCTGAAGACCCCAAGATCTCATGTTAAGTGAGAAAGGTTTGCAAAC 2400  
Db 2341 CTAAATTAATGCTGAAGACCCCAAGATCTCATGTTAAGTGAGAAAGGTTTGCAAAC 2400  
OY 2401 AAAGATCTGAGAGATGACATATTTCAATGCTGCTGCTGATGATGACACTCAGG 2460  
Db 2401 AAAGATCTGAGAGATGACATATTTCAATGCTGCTGCTGATGATGACACTCAGG 2460  
OY 2461 AAAGTATCTGTTACTGGAAGTTAGCACTCTAGGGAAGGCAAAAACAGAACCAATAAT 2520  
Db 2461 AAAGTATCTGTTACTGGAAGTTAGCACTCTAGGGAAGGCAAAAACAGAACCAATAAT 2520  
OY 2521 GTGTAGTCAAGTGTCTAGCATTGTGAAGCCCAAGGCACTAATGATGTTTCCAAG 2580  
Db 2521 GTGTAGTCAAGTGTCTAGCATTGTGAAGCCCAAGGCACTAATGATGTTTCCAAG 2580  
OY 2581 ATTAATGAAGATGACACAGAGGCTTTAAGTATCCATTGGGACATGAGTTAACCACTC 2640  
Db 2581 ATTAATGAAGATGACACAGAGGCTTTAAGTATCCATTGGGACATGAGTTAACCACTC 2640  
OY 2641 GGGAAACAAGCATAGAAATGGAAGAGTGAACCTGATGCTCAGTATTTGCAAGATACAT 2700  
Db 2641 GGGAAACAAGCATAGAAATGGAAGAGTGAACCTGATGCTCAGTATTTGCAAGATACAT 2700  
OY 2701 TCAAGTTTCAAGGCGCAGATCATTTGCTGTTTCAAAATCCAGAAATGCGAAGAGG 2760  
Db 2701 TCAAGTTTCAAGGCGCAGATCATTTGCTGTTTCAAAATCCAGAAATGCGAAGAGG 2760

```

OY 2761 AATGTCCAACTTCTCTGCCACTGTGGGTCCTTAAAGAAACAAAGTCCAAAGTCACTT 2820
    |||||
Db 2761 AATGTCCAACTTCTCTGCCACTGTGGGTCCTTAAAGAAACAAAGTCCAAAGTCACTT 2820
OY 2821 TTGAATGTGAACAAAGGAAGAAATTCAGAGAAAGATGAGTCTTAATATCAAGCTGTAC 2880
    |||||
Db 2821 TTGAATGTGAACAAAGGAAGAAATTCAGAGAAAGATGAGTCTTAATATCAAGCTGTAC 2880
OY 2881 AGACAGTAAATATCACTGACAGGCTTTCGTGGTGTGTCAGAAAGATAGACAGTGTGATA 2940
    |||||
Db 2881 AGACAGTAAATATCACTGACAGGCTTTCGTGGTGTGTCAGAAAGATAGACAGTGTGATA 2940
OY 2941 ATGCCAAATGTACTATCAAGAGAGGCTCTAGGTTTGTCTATCTCACTCACTTCAGAGCA 3000
    |||||
Db 2941 ATGCCAAATGTACTATCAAGAGAGGCTCTAGGTTTGTCTATCTCACTCACTTCAGAGCA 3000
OY 3001 AGCAAACTGGACATCTTACCTCCAAATTAACATGACCTTTTACAAACCCTATCTGTAC 3060
    |||||
Db 3001 AGCAAACTGGACATCTTACCTCCAAATTAACATGACCTTTTACAAACCCTATCTGTAC 3060
OY 3061 CACCACTTTTCCATCAAGTCAATTTGTTTAAACTAATGTAGAAATAATCTGTAGAG 3120
    |||||
Db 3061 CACCACTTTTCCATCAAGTCAATTTGTTTAAACTAATGTAGAAATAATCTGTAGAG 3120
OY 3121 AAAACTTTGAGAACATTCATGTCACCTGAGAAAGAAATGGAAATGAGAACATTTGCAA 3180
    |||||
Db 3121 AAAACTTTGAGAACATTCATGTCACCTGAGAAAGAAATGGAAATGAGAACATTTGCAA 3180
OY 3181 GTACAGTGTGACACAAATTAGCCGTAATACATTTAGAGAAATGTTTAAAGAGCCAGCT 3240
    |||||
Db 3181 GTACAGTGTGACACAAATTAGCCGTAATACATTTAGAGAAATGTTTAAAGAGCCAGCT 3240
OY 3241 CAAGCAATATTAATGAAGTAGGTTCCAGTACTAATGAAGTGGGCTCCAGTATTAAATGAA 3300
    |||||
Db 3241 CAAGCAATATTAATGAAGTAGGTTCCAGTACTAATGAAGTGGGCTCCAGTATTAAATGAA 3300
OY 3301 TTAGTTCCTCCAGTATGAGAAACATTCACAGCAGAACTAGTACAAACAGAGGCCCAAAATGGA 3360
    |||||
Db 3301 TTAGTTCCTCCAGTATGAGAAACATTCACAGCAGAACTAGTACAAACAGAGGCCCAAAATGGA 3360
OY 3361 ATGCTATGCTAGATAGAGGGTTTTCGCAACCTGAGGCTCTAATAACAAAGTCTTCCTGGA 3420
    |||||
Db 3361 ATGCTATGCTAGATAGAGGGTTTTCGCAACCTGAGGCTCTAATAACAAAGTCTTCCTGGA 3420
OY 3421 GTAATTTGAAGCATCTCTGAATAAAAAAGCAGAAATATGAAGAAGTAGTTCAAGACTGTA 3480
    |||||
Db 3421 GTAATTTGAAGCATCTCTGAATAAAAAAGCAGAAATATGAAGAAGTAGTTCAAGACTGTA 3480
OY 3481 ATACAGATTTCTCTCCATATCTGATTTTCAGATACTTAGAACAGCCATATGGGAAGTAGTC 3540
    |||||
Db 3481 ATACAGATTTCTCTCCATATCTGATTTTCAGATACTTAGAACAGCCATATGGGAAGTAGTC 3540
OY 3541 ATGCATCTCAGGTTTGTCTGAGACACCTGATGACCTGTAGATGATGATGGAATAAAG 3600
    |||||
Db 3541 ATGCATCTCAGGTTTGTCTGAGACACCTGATGACCTGTAGATGATGATGGAATAAAG 3600
OY 3601 AAGATCTAGTTTGTCTGAATAATGACATTAAGGAAAGTCTGCTTTTAAAGCAAAAGCG 3660
    |||||
Db 3601 AAGATCTAGTTTGTCTGAATAATGACATTAAGGAAAGTCTGCTTTTAAAGCAAAAGCG 3660
OY 3661 TTCAGAAAGGAGAGCTTAGCAGAGAGTCCAGCCCTTACACCCATACACATTTGGCTCAG 3720
    |||||
Db 3661 TTCAGAAAGGAGAGCTTAGCAGAGAGTCCAGCCCTTACACCCATACACATTTGGCTCAG 3720
OY 3721 GTTACCGAAGAGGGGCCAGAAATTTAGAGTCTCAGAAAGAACTTATCTAGTGAAGATG 3780
    |||||
Db 3721 GTTACCGAAGAGGGGCCAGAAATTTAGAGTCTCAGAAAGAACTTATCTAGTGAAGATG 3780
OY 3781 AAGAGCTTCCCTGCTTCCAAACCTTGTATTGGTAAAGTAAACAATATACCTTCTCAGT 3840
    |||||
Db 3781 AAGAGCTTCCCTGCTTCCAAACCTTGTATTGGTAAAGTAAACAATATACCTTCTCAGT 3840
OY 3841 CTACTAGGCATAGACACCGTTCCTACCGAGTGTCTGTCTAAGAACACAGAGAGAAATTTAT 3900
    |||||
    |||||
Db 3841 CTACTAGGCATAGACACCGTTCCTACCGAGTGTCTGTCTAAGAACACAGAGAGAAATTTAT 3900
    |||||
OY 3901 TATCATTTGAAGAAATAGCTTAAATGACTGCGATGTAACCAAGTAATATTTGGCAAGGCATCTC 3960
    |||||
Db 3901 TATCATTTGAAGAAATAGCTTAAATGACTGCGATGTAACCAAGTAATATTTGGCAAGGCATCTC 3960
OY 3961 AGGAACATCTACCTTAGAGAGAAACAAATGTTCTGTAGTACTGTTTCTTTCACACTGCA 4020
    |||||
Db 3961 AGGAACATCTACCTTAGAGAGAAACAAATGTTCTGTAGTACTGTTTCTTTCACACTGCA 4020
OY 4021 GTGAATTTGAAGAGCTTGACTGCAAAATCAAAACACCCAGATCTCTTCTGATTTGTTCTT 4080
    |||||
Db 4021 GTGAATTTGAAGAGCTTGACTGCAAAATCAAAACACCCAGATCTCTTCTGATTTGTTCTT 4080
OY 4021 GTGAATTTGAAGAGCTTGACTGCAAAATCAAAACACCCAGATCTCTTCTGATTTGTTCTT 4080
    |||||
Db 4021 GTGAATTTGAAGAGCTTGACTGCAAAATCAAAACACCCAGATCTCTTCTGATTTGTTCTT 4080
OY 4081 CCAACCAATGAGGATCTCACTGCAAGCCAGGAGTGTGTGATGATGCAAGAAATTTG 4140
    |||||
Db 4081 CCAACCAATGAGGATCTCACTGCAAGCCAGGAGTGTGTGATGATGCAAGAAATTTG 4140
OY 4141 TTTTCAGATGATGAAAGAAAGAGAGAGGCGCTTGAAGAAATATATCAAGAGAGCAAGCA 4200
    |||||
Db 4141 TTTTCAGATGATGAAAGAAAGAGAGAGGCGCTTGAAGAAATATATCAAGAGAGCAAGCA 4200
OY 4201 TGAATTTCAAACTTAGGTTGAAGAGCAGATCTGGGTTGTGAGAGTGAACACGCTCTGGA 4260
    |||||
Db 4201 TGAATTTCAAACTTAGGTTGAAGAGCAGATCTGGGTTGTGAGAGTGAACACGCTCTGGA 4260
OY 4261 ACTGCTCAGGCGTATCTCTCAGAGTACATTTTAACTTAACTCAGCAGAGAGATACCATGC 4320
    |||||
Db 4261 ACTGCTCAGGCGTATCTCTCAGAGTACATTTTAACTTAACTCAGCAGAGAGATACCATGC 4320
OY 4321 AACATAACTGTATTAAGCTCCAGCAGAGAAATGGCTGACATAGAGCTGTGTGAAGACAGC 4380
    |||||
Db 4321 AACATAACTGTATTAAGCTCCAGCAGAGAAATGGCTGACATAGAGCTGTGTGAAGACAGC 4380
OY 4381 ATGGGAGCCAGCCTTCTAAGAGTACCTTCCATCATAGTACTCTTGTGCTTGGCTTGA 4440
    |||||
Db 4381 ATGGGAGCCAGCCTTCTAAGAGTACCTTCCATCATAGTACTCTTGTGCTTGGCTTGA 4440
OY 4441 ACCTGCGCAATTCAGAAACAAAGCAGATCAGAAACACAGATTTAACTTACAGAAAGTA 4500
    |||||
Db 4441 ACCTGCGCAATTCAGAAACAAAGCAGATCAGAAACACAGATTTAACTTACAGAAAGTA 4500
OY 4501 GTGAATTAACCTTATAGCCAGCAATTCAGAAAGCCCTTCTGTGCAAGTTTGAAGTGTG 4560
    |||||
Db 4501 GTGAATTAACCTTATAGCCAGCAATTCAGAAAGCCCTTCTGTGCAAGTTTGAAGTGTG 4560
OY 4561 CAGATAGTCTTACCAGTAAAAAATTAAGAACACAGAGAGTGAAGAGTATCCCTTCTAAT 4620
    |||||
Db 4561 CAGATAGTCTTACCAGTAAAAAATTAAGAACACAGAGAGTGAAGAGTATCCCTTCTAAT 4620
OY 4621 GCCCATCATATGATGATAGTGTGTACATGACAGAGTGTCTGGAGTCTTCAGAAATAGA 4680
    |||||
Db 4621 GCCCATCATATGATGATAGTGTGTACATGACAGAGTGTCTGGAGTCTTCAGAAATAGA 4680
OY 4681 ACTACCATCTCAAGAGAGAGAGTCAATTAAGGTTGTGATGTGAGAGAGCAACAGCTGGA 4740
    |||||
Db 4681 ACTACCATCTCAAGAGAGAGAGTCAATTAAGGTTGTGATGTGAGAGAGCAACAGCTGGA 4740
OY 4741 AGTCTGGGCGCACAGATTTGAGAGGAACATCTTACTTGGCAAGGCAAGATCTGAGGGA 4800
    |||||
Db 4741 AGTCTGGGCGCACAGATTTGAGAGGAACATCTTACTTGGCAAGGCAAGATCTGAGGGA 4800
OY 4801 CCCCTTACCTGGAATGTGAATCAGAGCTTCTCTGATATACCTGAAATGTATCTCTG 4860
    |||||
Db 4801 CCCCTTACCTGGAATGTGAATCAGAGCTTCTCTGATATACCTGAAATGTATCTCTG 4860
OY 4861 AAGACAGAGAGAGAGAGAGTGTGTTGGCAACATACATCTTCAACCTTGTGATTTGA 4920
    |||||
Db 4861 AAGACAGAGAGAGAGAGAGTGTGTTGGCAACATACATCTTCAACCTTGTGATTTGA 4920
OY 4921 AAGTTCCTCCAAATTTGAAGTTGCAAGATCTGCCAGAGTCCAGCTGCTCATACTACTG 4980
    |||||
    |||||

```

Db 4921 AAGTCCCAATGTGAAGTGCAGATCTGCCAGAGTCCAGCTGCTCTACTACTG 4980  
 QY 4981 ATACTGCTGGGTATTAATGCAATGGAAGAAAGTGTGAGCAGGAGAGAACCCAGAAATTGACAG 5040  
 Db 4981 ATACGCTGGGTATTAATGCAATGGAAGAAAGTGTGAGCAGGAGAGAACCCAGAAATTGACAG 5040  
 QY 5041 CTTCACAGAAAGGCTCAACAAAGATGTCATGCTGTGCTGCGCTGACCCAGAG 5100  
 Db 5041 CTTCACAGAAAGGCTCAACAAAGATGTCATGCTGTGCTGCGCTGACCCAGAG 5100  
 QY 5101 AATTATCTGCTGACAAAGTGTGACAAAGAACCAATCATCTTAATCTAATTA 5160  
 Db 5101 AATTATCTGCTGACAAAGTGTGACAAAGAACCAATCATCTTAATCTAATTA 5160  
 QY 5161 CTGAAGAGACTACTCATCTGTTATGAAGACAGATGCTGAGTTGTGTGAGAGGAGAC 5220  
 Db 5161 CTGAAGAGACTACTCATCTGTTATGAAGACAGATGCTGAGTTGTGTGAGAGGAGAC 5220  
 QY 5221 TGAATATTTCTAGGAATTCGGGAGAGAAATGGGTAGTATTTCTGCTGAGCC 5280  
 Db 5221 TGAATATTTCTAGGAATTCGGGAGAGAAATGGGTAGTATTTCTGCTGAGCC 5280  
 QY 5281 AGCTATTAAGAAAGAAATGCTGATGAGCATGATTTGAACTGAGAGAGATGCG 5340  
 Db 5281 AGCTATTAAGAAAGAAATGCTGATGAGCATGATTTGAACTGAGAGAGATGCG 5340  
 QY 5341 TCAATGGAAGAACCCAGAGGTCGAAGCGAGAGAAATCCAGAGAGAAAGATCT 5400  
 Db 5341 TCAATGGAAGAACCCAGAGGTCGAAGCGAGAGAAATCCAGAGAGAAAGATCT 5400  
 QY 5401 TCAGGGGCTAGAAATCTGTTGCTATGGGCCCTTCACCAAGATGCCAGATCAACTG 5460  
 Db 5401 TCAGGGGCTAGAAATCTGTTGCTATGGGCCCTTCACCAAGATGCCAGATCAACTG 5460  
 QY 5461 AATGATGATGACAGTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 5520  
 Db 5461 AATGATGATGACAGTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 5520  
 QY 5521 GCACAGGTGTCACCAATTTGTTGTTGTCAGCCAGATGCTGAGCAGAGAGACATGCT 5580  
 Db 5521 GCACAGGTGTCACCAATTTGTTGTTGTCAGCCAGATGCTGAGCAGAGAGACATGCT 5580  
 QY 5581 TCATGCAATTTGGGCGAGATGTGTGAGGCACTGTGTGAGCCGAGAGATGGTGTGACA 5640  
 Db 5581 TCATGCAATTTGGGCGAGATGTGTGAGGCACTGTGTGAGCCGAGAGATGGTGTGACA 5640  
 QY 5641 GTGAGACCTCTACAGAGTGCAGAGAGCTGAGACCTACTGATACCCAGATCCGCCACA 5700  
 Db 5641 GTGAGACCTCTACAGAGTGCAGAGAGCTGAGACCTACTGATACCCAGATCCGCCACA 5700  
 QY 5701 GCCACTACTGA 5711  
 Db 5701 GCCACTACTGA 5711  
 RESULT 14  
 LOCUS 181034 5914 bp DNA Linear PAT 10-JUN-1998  
 DEFINITION Sequence 1 from patent US 5710001.  
 ACCESSION 181034  
 VERSION 181034.1 GI:3209324  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE Unclassified.  
 1 (bases 1 to 5914)  
 AUTHORS Skolnick, M.H., Goldgar, D.E., Miki, Y., Swenson, J., Kamb, A.,  
 and Futreal, P. Andrew.  
 TITLE 17q-linked breast and ovarian cancer susceptibility gene  
 JOURNAL Patent: US 5710001-A 1 20-JAN-1998;  
 FEATURES Location/Qualifiers  
 source 1..5914

BASE COUNT 2006 a 1156 c 1316 g 1436 t  
 ORIGIN  
 Query Match 99.9%; Score 5707.8; DB 6; Length 5914;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 5709; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 AGCTGCTGAGACCTCTGAGACCCCGACAGGCTGTGGGTTCTCAGATACTGGCC 60  
 Db 1 AGCTGCTGAGACCTCTGAGACCCCGACAGGCTGTGGGTTCTCAGATACTGGCC 60  
 QY 61 CCTGGCTCAGAGGCCCTTCAACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
 Db 61 CCTGGCTCAGAGGCCCTTCAACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
 QY 121 TGGATTTATCTGCTCTTGGCGTTGAGAGATGCAAAATGCTATTAATCTATGCAAGAA 180  
 Db 121 TGGATTTATCTGCTCTTGGCGTTGAGAGATGCAAAATGCTATTAATCTATGCAAGAA 180  
 QY 181 TCTTAGAGTGGCCATCTGCTGAGTGTGATCAAGAAAGCTGTCCCAAGAGTGTGACC 240  
 Db 181 TCTTAGAGTGGCCATCTGCTGAGTGTGATCAAGAAAGCTGTCCCAAGAGTGTGACC 240  
 QY 241 ACATATTTTGGCAATTTTGCATGCTGCAACTTCTCAACAGAGAAAGGCTTCCACAGT 300  
 Db 241 ACATATTTTGGCAATTTTGCATGCTGCAACTTCTCAACAGAGAAAGGCTTCCACAGT 300  
 QY 301 GTCTTTATGTAAGATGATATTAACCAAGAGGCTTCAAGAAAGTTCAGATTTAGTC 360  
 Db 301 GTCTTTATGTAAGATGATATTAACCAAGAGGCTTCAAGAAAGTTCAGATTTAGTC 360  
 QY 361 AACCTGTTGAAGAGCTATGAAATCATTTGCTTTCAGCTTGCAGTTCAGAGTTGGAGT 420  
 Db 361 AACCTGTTGAAGAGCTATGAAATCATTTGCTTTCAGCTTGCAGTTCAGAGTTGGAGT 420  
 QY 421 ATGCAAGAGCTATTAATTTTCCAAAAAGAAATTAATCTCTGCAATCTTAAGATG 480  
 Db 421 ATGCAAGAGCTATTAATTTTCCAAAAAGAAATTAATCTCTGCAATCTTAAGATG 480  
 QY 481 AAGTTTCTATCTCAAAAGTATGGCTTACAAACCGTCCAAAAAGCTTCTACAGAGTG 540  
 Db 481 AAGTTTCTATCTCAAAAGTATGGCTTACAAACCGTCCAAAAAGCTTCTACAGAGTG 540  
 QY 541 AACCGAAATCTCTCTGAGAGAACCACTGACAGTCCACTCTTAACCTGGAA 600  
 Db 541 AACCGAAATCTCTCTGAGAGAACCACTGACAGTCCACTCTTAACCTGGAA 600  
 QY 601 CTGTGAGAACTCTGAGAGCAAGCAGGATACCAACTCAAAAGACGTCGTCTACATG 660  
 Db 601 CTGTGAGAACTCTGAGAGCAAGCAGGATACCAACTCAAAAGACGTCGTCTACATG 660  
 QY 661 AATTGGATCTGATTTCTTCAAGATACCGTTAATTAAGCACTTATTTGCACTGGAG 720  
 Db 661 AATTGGATCTGATTTCTTCAAGATACCGTTAATTAAGCACTTATTTGCACTGGAG 720  
 QY 721 ATCAAGATTTGTTCAAAATCAACCCCTCAAGAACCGAGATGAATCACTGTTGGATCTG 780  
 Db 721 ATCAAGATTTGTTCAAAATCAACCCCTCAAGAACCGAGATGAATCACTGTTGGATCTG 780  
 QY 781 CAAAAAGGCTGCTGTGAAATTTTCTGAGAGGATGTAACAAATCTGAAACATCAAC 840  
 Db 781 CAAAAAGGCTGCTGTGAAATTTTCTGAGAGGATGTAACAAATCTGAAACATCAAC 840  
 QY 841 CCAATTAATGATTTGAGACCACTGAGAGGCTGCAAGCTGCAAGAGCATCAAGAAAGT 900  
 Db 841 CCAATTAATGATTTGAGACCACTGAGAGGCTGCAAGCTGCAAGAGCATCAAGAAAGT 900  
 QY 901 ATCAGGAGATGCTGTTTCAAACTGCTGAGAGGATGAGAGGATGAGAGGATGAGAGG 960  
 Db 901 ATCAGGAGATGCTGTTTCAAACTGCTGAGAGGATGAGAGGATGAGAGGATGAGAGG 960  
 QY 961 GCTCATTACAGCATGAGAGAGAGGATTTATTACTACTAAAGACGATGAATGTAGAAA 1020

|||||  
Db 961 GCTCATTCACGACATGAGAACAGCAGTTATTACTACTAAGACAGATGATGTAGAAA 1020  
QY 1021 AGGCTAATTCCTGTAATAAAAAGCAACACCTGGCTTTAGCAAGAGCCAACTAATACAGAT 1080  
Db 1021 AGGCTAATTCCTGTAATAAAAAGCAACACCTGGCTTTAGCAAGAGCCAACTAATACAGAT 1080  
QY 1081 GGGCTGGAAGTAAAGGAACATGTAATAGTGGCGACTCCGACGACAGAAAAAAGGTAG 1140  
Db 1081 GGGCTGGAAGTAAAGGAACATGTAATAGTGGCGACTCCGACGACAGAAAAAAGGTAG 1140  
QY 1141 ATCTGAATGCTGATCCCTGTGTGAGAGAAAAAGTAATGGAATAGCAGAACTGCCATGCT 1200  
Db 1141 ATCTGAATGCTGATCCCTGTGTGAGAGAAAAAGTAATGGAATAGCAGAACTGCCATGCT 1200  
QY 1201 CAGAGATCCTAGAGTACTGAAAGATGTTCTTGATTAACACTAAATGAGCAATTCGA 1260  
Db 1201 CAGAGATCCTAGAGTACTGAAAGATGTTCTTGATTAACACTAAATGAGCAATTCGA 1260  
QY 1261 AAGTTAATGAGTGGTTTTCCAGAAAGTATGAACTGTTAGTTCTGATGACTCAGATGATG 1320  
Db 1261 AAGTTAATGAGTGGTTTTCCAGAAAGTATGAACTGTTAGTTCTGATGACTCAGATGATG 1320  
QY 1321 GGGAGTCTGATCAATCAATGCCAAAGTACCTGATGTAATGGACGTTCTAATGAGTAGATG 1380  
Db 1321 GGGAGTCTGATCAATCAATGCCAAAGTACCTGATGTAATGGACGTTCTAATGAGTAGATG 1380  
QY 1381 AATATTCTGTTCTTCAGAGAAATATAGACTAGCGGAGATGATGCTCATGAGGCTTTAA 1440  
Db 1381 AATATTCTGTTCTTCAGAGAAATATAGACTAGCGGAGATGATGCTCATGAGGCTTTAA 1440  
QY 1441 TATGTAAAGTGAAGAAGTTCACCTCAATCAGTAGAGAGTAATATTGAACACAAAATAT 1500  
Db 1441 TATGTAAAGTGAAGAAGTTCACCTCAATCAGTAGAGAGTAATATTGAACACAAAATAT 1500  
QY 1501 TTGGGAAAACCTATCGGAAAGAGCAAGCCTCCCAACTTAAAGCATGTAACTGAAAATC 1560  
Db 1501 TTGGGAAAACCTATCGGAAAGAGCAAGCCTCCCAACTTAAAGCATGTAACTGAAAATC 1560  
QY 1561 TAATTTATGAGAGCATTTGTTACTAGGACACAGATTAATACAGAGCGTCCCTCAANA 1620  
Db 1561 TAATTTATGAGAGCATTTGTTACTAGGACACAGATTAATACAGAGCGTCCCTCAANA 1620  
QY 1621 AATTTAAAGCGTAAAGGAGACCTATACATCAGGCTTTCCTGAGAGATTTTATCAAGAA 1680  
Db 1621 AATTTAAAGCGTAAAGGAGACCTATACATCAGGCTTTCCTGAGAGATTTTATCAAGAA 1680  
QY 1681 CAGATTTGGCAGTTCAAAAAGACTCTGTAATGATAAATCAGGGAACTAACCAAGCGAGC 1740  
Db 1681 CAGATTTGGCAGTTCAAAAAGACTCTGTAATGATAAATCAGGGAACTAACCAAGCGAGC 1740  
QY 1741 AGAATGGTCAAGTGAATGATTTACTAATAGTGGTCAATGAGAAATTAACAAAGGTGAT 1800  
Db 1741 AGAATGGTCAAGTGAATGATTTACTAATAGTGGTCAATGAGAAATTAACAAAGGTGAT 1800  
QY 1801 CTATTCAGATGAGAAAAATCTTAACCAATAGAACTCTGAAAAAAGAAATCTGCTTCA 1860  
Db 1801 CTATTCAGATGAGAAAAATCTTAACCAATAGAACTCTGAAAAAAGAAATCTGCTTCA 1860  
QY 1861 AAACGAAAGCTGAACCTATTAAGCAGAGTAAAGCAATATGAACTGAAATTAATATCC 1920  
Db 1861 AAACGAAAGCTGAACCTATTAAGCAGAGTAAAGCAATATGAACTGAAATTAATATCC 1920  
QY 1921 ACAATTCAAAAGACCTAAAGAAATATGCTGAGAGAGAGTCTTCTACAGAGCATATTC 1980  
Db 1921 ACAATTCAAAAGACCTAAAGAAATATGCTGAGAGAGAGTCTTCTACAGAGCATATTC 1980  
QY 1981 ATGCGCTTAAGTAACTAGTACAGTAAATCTAAGCCACCTAATTTGATGAAATTCGAA 2040  
Db 1981 ATGCGCTTAAGTAACTAGTACAGTAAATCTAAGCCACCTAATTTGATGAAATTCGAA 2040  
QY 2041 TTGATAGTTGTTTACGAGTGAAGAGATTAAGAAAAAGTACACCAATGCGCAGTCA 2100  
Db 2041 TTGATAGTTGTTTACGAGTGAAGAGATTAAGAAAAAGTACACCAATGCGCAGTCA 2100

Db 2041 TTGATAGTTGTTTACGAGTGAAGAGATTAAGAAAAAGTACACCAATGCGCAGTCA 2100  
QY 2101 GGCACAGAGAAACCTCAACTCATGTAAGTAAAGAACTGCAACTGGAGCCAAAGAA 2160  
Db 2101 GGCACAGAGAAACCTCAACTCATGTAAGTAAAGAACTGCAACTGGAGCCAAAGAA 2160  
QY 2161 GTAAACAGCCAAATGAAGACAGCAAGTAAAGACATGACATGATCTTCCAGACTGA 2220  
Db 2161 GTAAACAGCCAAATGAAGACAGCAAGTAAAGACATGACATGATCTTCCAGACTGA 2220  
QY 2221 AGTTAACAATATGACCTGCTTTCTTACTAAGTGTTCAAATACCAAGTAAAGAAAT 2280  
Db 2221 AGTTAACAATATGACCTGCTTTCTTACTAAGTGTTCAAATACCAAGTAAAGAAAT 2280  
QY 2281 TTGTCATTCCTGAGCTTCCAAAGAGAAAGAAAGAAAGCAATGAAACGTTAAAGTCT 2340  
Db 2281 TTGTCATTCCTGAGCTTCCAAAGAGAAAGAAAGAAAGCAATGAAACGTTAAAGTCT 2340  
QY 2341 CTAATTAATGCTGAAGACCCCAAGATCTCATGTTAAGTGAAGAAAGGTTTCCAAAGT 2400  
Db 2341 CTAATTAATGCTGAAGACCCCAAGATCTCATGTTAAGTGAAGAAAGGTTTCCAAAGT 2400  
QY 2401 AAAGATCTGTAGAGAGTACAGATATTCATTTGCTGACTGATGATGATGATGATG 2460  
Db 2401 AAAGATCTGTAGAGAGTACAGATATTCATTTGCTGACTGATGATGATGATGATG 2460  
QY 2461 AAAGATCTGTAGAGAGTACAGATATTCATTTGCTGACTGATGATGATGATGATG 2520  
Db 2461 AAAGATCTGTAGAGAGTACAGATATTCATTTGCTGACTGATGATGATGATGATG 2520  
QY 2521 GTGTGAGTCAAGTGTGACAGATTTTGAAGAAACCCCAAGGAGTAAATTCATGTTTCCAA 2580  
Db 2521 GTGTGAGTCAAGTGTGACAGATTTTGAAGAAACCCCAAGGAGTAAATTCATGTTTCCAA 2580  
QY 2581 ATATTAAGAAATGACACGAAAGGCTTTAAGTATCATTTGGACATGAAAGTAAACACAG 2640  
Db 2581 ATATTAAGAAATGACACGAAAGGCTTTAAGTATCATTTGGACATGAAAGTAAACACAG 2640  
QY 2641 GGGAAACAGCATGTAAGTGAAGAAAGTGAACCTGATGCTGATGCTGATGCTGATGCT 2700  
Db 2641 GGGAAACAGCATGTAAGTGAAGAAAGTGAACCTGATGCTGATGCTGATGCTGATGCT 2700  
QY 2701 TCAAGGTTTCAAGCGGACATCTTGTCTGTTTCAATCCAGAGAAATGCAAGAGAG 2760  
Db 2701 TCAAGGTTTCAAGCGGACATCTTGTCTGTTTCAATCCAGAGAAATGCAAGAGAG 2760  
QY 2761 AATGTGCAACATCTCTGCCCCACTCTGCGTCTTAAAGAAACAAATGCTCACTT 2820  
Db 2761 AATGTGCAACATCTCTGCCCCACTCTGCGTCTTAAAGAAACAAATGCTCACTT 2820  
QY 2821 TTGAAATGTAACAAAGAAAGAAATCAAGAAAGTAAAGTAAATGATCAAGGCTGTAC 2880  
Db 2821 TTGAAATGTAACAAAGAAAGAAATCAAGAAAGTAAAGTAAATGATCAAGGCTGTAC 2880  
QY 2881 AGACAGTTAATATCACTGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2940  
Db 2881 AGACAGTTAATATCACTGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2940  
QY 2941 ATGCCAAATATGATATCAAGAGAGCTGAGGTTTGTCTATCANTCTAGTTCAAGAGCA 3000  
Db 2941 ATGCCAAATATGATATCAAGAGAGCTGAGGTTTGTCTATCANTCTAGTTCAAGAGCA 3000  
QY 3001 ACGAACCTGACATCTTACCTCCAAATTAACATGCACTTTTACAAAACCATATGCTATAC 3060  
Db 3001 ACGAACCTGACATCTTACCTCCAAATTAACATGCACTTTTACAAAACCATATGCTATAC 3060  
QY 3061 CACCACTTTTCCATCAAGTCAATTTGTTTAAACTAAATGTAAGAAAAATCTGCTAGAG 3120  
Db 3061 CACCACTTTTCCATCAAGTCAATTTGTTTAAACTAAATGTAAGAAAAATCTGCTAGAG 3120  
QY 3121 AAAACTTTGAGAACATTTCAATGTCACTGAAAGAGAAATGGAATGGAACATTTCCAA 3180  
Db 3121 AAAACTTTGAGAACATTTCAATGTCACTGAAAGAGAAATGGAATGGAACATTTCCAA 3180

```

OY 3181 GTACAGTGACACAAATTTAGCCGTAAATTAACATTAGAGAAAATGTTTTTAAGAGCCACT 3240
    |||||||
Db 3181 GTACAGTGACACAAATTTAGCCGTAAATTAACATTAGAGAAAATGTTTTTAAGAGCCACT 3240
OY 3241 CAAGCAATATTAATAGTAGTGGTTCAGTACTAATGAAGTGGCTCCAGTATTAATGA 3300
    |||||||
Db 3241 CAAGCAATATTAATAGTAGTGGTTCAGTACTAATGAAGTGGCTCCAGTATTAATGA 3300
OY 3301 TAGGTTCCAGTGATGATAAACAATTCAGAGAACTAGTAGAAACAGAGAGGCCAAATTTGA 3360
    |||||||
Db 3301 TAGGTTCCAGTGATGATAAACAATTCAGAGAACTAGTAGAAACAGAGAGGCCAAATTTGA 3360
OY 3361 ATGCTATGCTTAGATTAGGGGTTTTGCAACCTGAGAGTCTATTAACAAAGTCTTCTGGA 3420
    |||||||
Db 3361 ATGCTATGCTTAGATTAGGGGTTTTGCAACCTGAGAGTCTATTAACAAAGTCTTCTGGA 3420
OY 3421 GTAATTGTAAGCATCTCTGAATAAATAAAGCAAGATATGAAGAACTAGTCAAGCTGTTA 3480
    |||||||
Db 3421 GTAATTGTAAGCATCTCTGAATAAATAAAGCAAGATATGAAGAACTAGTCAAGCTGTTA 3480
OY 3481 ATACGATTTCTCTCATATCTGATTTCAGATTACTTGAACAGCCTATGGAGTACTC 3540
    |||||||
Db 3481 ATACGATTTCTCTCATATCTGATTTCAGATTACTTGAACAGCCTATGGAGTACTC 3540
OY 3541 ATGCATCTCAGGTTGTTCTGAGACACCTGATGACCTGTAGATGATGTAATAAAG 3600
    |||||||
Db 3541 ATGCATCTCAGGTTGTTCTGAGACACCTGATGACCTGTAGATGATGTAATAAAG 3600
OY 3601 AAGATACTAGTTTCTGTAATAATGACATTAAGAAAGTTCGTGTTTTAGCAAAACCG 3660
    |||||||
Db 3601 AAGATACTAGTTTCTGTAATAATGACATTAAGAAAGTTCGTGTTTTAGCAAAACCG 3660
OY 3661 TCCAGAAAGGAGGCTTAGCAGAGTCTAGCCCTTACCCATTCACATTTGGTCCAG 3720
    |||||||
Db 3661 TCCAGAAAGGAGGCTTAGCAGAGTCTAGCCCTTACCCATTCACATTTGGTCCAG 3720
OY 3721 GTTACCGAAGAGGGGCCAAGAAATTAAGATCTCTCAGAGAACTTATCTAGTAGATG 3780
    |||||||
Db 3721 GTTACCGAAGAGGGGCCAAGAAATTAAGATCTCTCAGAGAACTTATCTAGTAGATG 3780
OY 3781 AAGAGCTTCCTGCTTCCACACCTGTTATTTGTTAAAGTAACAAATATACCTTCTCAGT 3840
    |||||||
Db 3781 AAGAGCTTCCTGCTTCCACACCTGTTATTTGTTAAAGTAACAAATATACCTTCTCAGT 3840
OY 3841 CTACTAGGCATAGACCGTGTCTACCGAGTGTCTGTCTAAGAACACAGAGAGATTTAT 3900
    |||||||
Db 3841 CTACTAGGCATAGACCGTGTCTACCGAGTGTCTGTCTAAGAACACAGAGAGATTTAT 3900
OY 3901 TATCATTTGAAAGAAATAGCTTAATGACTGCACTAACACAGTAATATTTGGCAAGCATCTC 3960
    |||||||
Db 3901 TATCATTTGAAAGAAATAGCTTAATGACTGCACTAACACAGTAATATTTGGCAAGCATCTC 3960
OY 3961 AGGAACATCACCTTAGTAGAGAAACAAATATCTGTACGCTTTCTTCAACAGTCA 4020
    |||||||
Db 3961 AGGAACATCACCTTAGTAGAGAAACAAATATCTGTACGCTTTCTTCAACAGTCA 4020
OY 4021 GTGATTTGAGAGACTTGACTCCAAATATCAAAACCCAGAGATCCTTCTGATGGTCTT 4080
    |||||||
Db 4021 GTGATTTGAGAGACTTGACTCCAAATATCAAAACCCAGAGATCCTTCTGATGGTCTT 4080
OY 4081 CCAAAACAATGAGGATCAGTCTGAAGCCAGGGAGTTGCTGCTAGTACAGAGATTTGG 4140
    |||||||
Db 4081 CCAAAACAATGAGGATCAGTCTGAAGCCAGGGAGTTGCTGCTAGTACAGAGATTTGG 4140
OY 4141 TTTTCAGATGATGAAGAAAGAGAAACGGCTTGAAGAAATATATCAGAGAGCAAAACA 4200
    |||||||
Db 4141 TTTTCAGATGATGAAGAAAGAGAAACGGCTTGAAGAAATATATCAGAGAGCAAAACA 4200
OY 4201 TGGATTTCAAACTTAGGTGAGAGCAGATCTGGGTGAGAGTGAAGCAAGCCTCTCTGAG 4260
    |||||||
Db 4201 TGGATTTCAAACTTAGGTGAGAGCAGATCTGGGTGAGAGTGAAGCAAGCCTCTCTGAG 4260

```

---

```

OY 4261 ACTGCTCAGGGCTATCTCTCAGAGTGAACATTTTAACCACTCAGCAGAGAGATACATGC 4320
    |||||||
Db 4261 ACTGCTCAGGGCTATCTCTCTCAGAGTGAACATTTTAACCACTCAGCAGAGAGATACATGC 4320
OY 4321 AACATTAACCTGATAAAGCTCCAGCAGAGAAATGGCTGAACCTAGAAAGCTGTGTTAGAACAGC 4380
    |||||||
Db 4321 AACATTAACCTGATAAAGCTCCAGCAGAGAAATGGCTGAACCTAGAAAGCTGTGTTAGAACAGC 4380
OY 4381 ATGGAGGCCACCTTCTTAACAGCTCCCTTCATCATATAAGTACCTTCTGCTTGGAG 4440
    |||||||
Db 4381 ATGGAGGCCACCTTCTTAACAGCTCCCTTCATCATATAAGTACCTTCTGCTTGGAG 4440
OY 4441 ACCTCGGAATTCAGAAACAAAGACATCAGAAAAAGAGCTATTAACTTCACAGAAAGTA 4500
    |||||||
Db 4441 ACCTCGGAATTCAGAAACAAAGACATCAGAAAAAGAGCTATTAACTTCACAGAAAGTA 4500
OY 4501 GTGAATFACCTATTAAGCCAGAAATCCAGAGCCCTTCTGCTGACAAAGTTGAGGTCTG 4560
    |||||||
Db 4501 GTGAATFACCTATTAAGCCAGAAATCCAGAGCCCTTCTGCTGACAAAGTTGAGGTCTG 4560
OY 4561 CAGATAGTTCTACAGTAATAAATAAAGAACAGAGTGAAGAGTCAATCCCTCTTAAT 4620
    |||||||
Db 4561 CAGATAGTTCTACAGTAATAAATAAAGAACAGAGTGAAGAGTCAATCCCTCTTAAT 4620
OY 4621 GCCCATCATTAAGATAGTGTGTACATGACAGTGTCTGAGAGTCTTCAGAAATAGAA 4680
    |||||||
Db 4621 GCCCATCATTAAGATAGTGTGTACATGACAGTGTCTGAGAGTCTTCAGAAATAGAA 4680
OY 4681 ACTACCATCTCAAGAGAGAGCTCATTAAGTTGTTGATGTGAGAGAGCAACAGCTGAG 4740
    |||||||
Db 4681 ACTACCATCTCAAGAGAGAGCTCATTAAGTTGTTGATGTGAGAGAGCAACAGCTGAG 4740
OY 4741 AGTCGGGCCACAGATTTGACGAAACATCTTACTTCCCAAGGCAAGATCTAGAGGAA 4800
    |||||||
Db 4741 AGTCGGGCCACAGATTTGACGAAACATCTTACTTCCCAAGGCAAGATCTAGAGGAA 4800
OY 4801 CCCCTTACCTGGAATCTGGAATCAGCCCTCTCTGATGAGCCCTGAAATCTGATCTCTG 4860
    |||||||
Db 4801 CCCCTTACCTGGAATCTGGAATCAGCCCTCTCTGATGAGCCCTGAAATCTGATCTCTG 4860
OY 4861 AAGACAGAGCCCCAGAGTCAAGTCTGTGGCAACATTCATCTTCAACCTCTCATTTGA 4920
    |||||||
Db 4861 AAGACAGAGCCCCAGAGTCAAGTCTGTGGCAACATTCATCTTCAACCTCTCATTTGA 4920
OY 4921 AAGTTCCTCCATTTGAAGTTCAGAAATCTGCCCCAGAGTCCAGTGTGCTCATCTACTG 4980
    |||||||
Db 4921 AAGTTCCTCCATTTGAAGTTCAGAAATCTGCCCCAGAGTCCAGTGTGCTCATCTACTG 4980
OY 4981 ATACTGCTGGGTATTAATGCAATGGAAGAAAGTGTGAGCAGGAGAAAGCCAGAAATTTGACAG 5040
    |||||||
Db 4981 ATACTGCTGGGTATTAATGCAATGGAAGAAAGTGTGAGCAGGAGAAAGCCAGAAATTTGACAG 5040
OY 5041 CTTCAACAGAAAGGCTCAACAAAGAAATGTCATGTTGTTGCTGTGAGCCCTGACCCAGAG 5100
    |||||||
Db 5041 CTTCAACAGAAAGGCTCAACAAAGAAATGTCATGTTGTTGCTGTGAGCCCTGACCCAGAG 5100
OY 5101 AATTTATGCTGCTGTACAGTTTCCAGAAACCCACATCTTAATCAATCAATTAATTA 5160
    |||||||
Db 5101 AATTTATGCTGCTGTACAGTTTCCAGAAACCCACATCTTAATCAATCAATTAATTA 5160
OY 5161 CTGAGAGAGACTACATCATGTTTATGAAGACAGATGCTGATTTGTGTGTAACAGGACAC 5220
    |||||||
Db 5161 CTGAGAGAGACTACATCATGTTTATGAAGACAGATGCTGATTTGTGTGTAACAGGACAC 5220
OY 5221 TGAATATTTTCTGAGAAATTCGAGAGGAAATGGGATAGTACTATTTCTGAGGTGACCC 5280
    |||||||
Db 5221 TGAATATTTTCTGAGAAATTCGAGAGGAAATGGGATAGTACTATTTCTGAGGTGACCC 5280
OY 5281 AGCTATTAAAGAAAGAAATGCTGAATGAGCATGATTTTGAAGTCAAGAGAAATGAG 5340
    |||||||
Db 5281 AGCTATTAAAGAAAGAAATGCTGAATGAGCATGATTTTGAAGTCAAGAGAAATGAG 5340
OY 5341 TCAATGAGAGAAACACCAAGCTCAAAAGCAGCAAGAAAGATCCAGAGACAGAAAGATCT 5400
    |||||||

```



```

Db      5341 TCAGTGGAGAAACCAACAGAGTCACAAAGGAGACAGAAATGCCAGAGACAGAAAGATCT 5400
QY      5401 TCAGGGGGGCTAGAAATCTGTGTCTATGAGGCCCTTCACCAACATGCCAGATCAACTGG 5460
Db      5401 TCAGGGGGGCTAGAAATCTGTGTCTATGAGGCCCTTCACCAACATGCCAGATCAACTGG 5460
QY      5461 AATGATGTGTACAGCTGTGTGTCTCTGTGTGAAGAGCTTTCATTCACCTG 5520
Db      5461 AATGATGTGTACAGCTGTGTGTCTCTGTGTGAAGAGCTTTCATTCACCTG 5520
QY      5521 GCACAGGTGTCCACCAATTTGTGTGTGACAGCCAGATGCTTGACAGAGACAGTGGCT 5580
Db      5521 GCACAGGTGTCCACCAATTTGTGTGTGACAGCCAGATGCTTGACAGAGACAGTGGCT 5580
QY      5581 TCCATGCAATTTGGGAGATGTGTGAGGACAGCTGTGTGACCCGAGAGTGGGTGTGACA 5640
Db      5581 TCCATGCAATTTGGGAGATGTGTGAGGACAGCTGTGTGACCCGAGAGTGGGTGTGACA 5640
QY      5641 GTGTAGCAGCTCTACAGAGTGCAGAGAGCTGACACCTCTGATACCCAGATCCCGACA 5700
Db      5641 GTGTAGCAGCTCTACAGAGTGCAGAGAGCTGACACCTCTGATACCCAGATCCCGACA 5700
QY      5641 GTGTAGCAGCTCTACAGAGTGCAGAGAGCTGACACCTCTGATACCCAGATCCCGACA 5700
Db      5641 GTGTAGCAGCTCTACAGAGTGCAGAGAGCTGACACCTCTGATACCCAGATCCCGACA 5700
QY      5701 GCCACTACTGA 5711
Db      5701 GCCACTACTGA 5711

RESULT 15
AR048660
LOCUS      AR048660 5711 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 4 from patent US 5821328.
ACCESSION  AR048660
VERSION     AR048660.1 GI:59711003
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 5711)
AUTHORS     King, M.-C., Friedman, L., Ostermeyer, B., Rowell, S., Lynch, E.,
            Szabo, C. and Lee, M.
            Genetic markers for breast, ovarian, and prostatic cancer
            Patent: US 5821328-A 4 13-Oct-1998;
            Location/Qualifiers
            source          1. 5711
                           /organism="Unknown"
BASE COUNT  1956 a 1099 c 1275 g 1381 t
ORIGIN
Query Match 99.9%; Score 5706.2; DB 6; Length 5711;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 5708; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      301 GTCCCTTATGTAGAAATGATATAACCAAAAGAGGCTTACAGAAAGTACGAGATTTAGTC 360
Db      301 GTCCCTTATGTAGAAATGATATAACCAAAAGAGGCTTACAGAAAGTACGAGATTTAGTC 360
QY      361 AACTTGTGAAGAGCTATTGAAAATCATTTTGTGCTTTTCAGCTTTGACACAGGTTTGAGT 420
Db      361 AACTTGTGAAGAGCTATTGAAAATCATTTTGTGCTTTTCAGCTTTGACACAGGTTTGAGT 420
QY      421 ATGCAACAGCTATTAATTTTGCAGAAAAGAAATTAACCTGCTGAAACATCTAAAAGATG 480
Db      421 ATGCAACAGCTATTAATTTTGCAGAAAAGAAATTAACCTGCTGAAACATCTAAAAGATG 480
QY      481 AAGTTTCTATTCATCCAAAGATATGGGCTACAGAAACCGTGCCAAAGACTTTTACAGATG 540
Db      481 AAGTTTCTATTCATCCAAAGATATGGGCTACAGAAACCGTGCCAAAGACTTTTACAGATG 540
QY      541 AACCCGAAATCCCTCTCTTGCAGAAAACGAGTGCAGTGTCCAACTCTTAACCTTGGA 600
Db      541 AACCCGAAATCCCTCTCTTGCAGAAAACGAGTGCAGTGTCCAACTCTTAACCTTGGA 600
QY      601 CTGTGAGAACTGTGAGACAAAGAGGAGATACACCTCAAAAGAGCTGTGTACATTTG 660
Db      601 CTGTGAGAACTGTGAGACAAAGAGGAGATACACCTCAAAAGAGCTGTGTACATTTG 660
QY      661 AATTGGATCTGATTTCTTGAAGATACCGTTAATAAGGCACTTAATGCAAGTGGAG 720
Db      661 AATTGGATCTGATTTCTTGAAGATACCGTTAATAAGGCACTTAATGCAAGTGGAG 720
QY      721 ATCAAGAAATTTTTCAAATACCCCTCAAGAACCGAGATGAAATCATGTTGATTCG 780
Db      721 ATCAAGAAATTTTTCAAATACCCCTCAAGAACCGAGATGAAATCATGTTGATTCG 780
QY      781 CAAAAGAGGCGCTTGTGAATTTTCTGAGAGCGGATGAACAAATPACTGAAATCATCAAC 840
Db      781 CAAAAGAGGCGCTTGTGAATTTTCTGAGAGCGGATGAACAAATPACTGAAATCATCAAC 840
QY      841 CCAGTAATTAATGATTTGAACACACCTGAGAGGCTGAGAGGATCCAGAAAAGT 900
Db      841 CCAGTAATTAATGATTTGAACACACCTGAGAGGCTGAGAGGATCCAGAAAAGT 900
QY      901 ATCAGGATAGTGTGTTTCAAACTTGCATGTGAGAGGCTGTGGCACAATATCTCATGCCA 960
Db      901 ATCAGGATAGTGTGTTTCAAACTTGCATGTGAGAGGCTGTGGCACAATATCTCATGCCA 960
QY      961 GCTCATTAACAGATGAGAACAGCAGTTTATTTACTCACTAAGACAGATGATGTGAAA 1020
Db      961 GCTCATTAACAGATGAGAACAGCAGTTTATTTACTCACTAAGACAGATGATGTGAAA 1020
QY      1021 AAGCTGAATTCGTATATAAAGCAACAGGCTGTAGCAGAGGCCAATACAGAT 1080
Db      1021 AAGCTGAATTCGTATATAAAGCAACAGGCTGTAGCAGAGGCCAATACAGAT 1080
QY      1081 GGGCTGGAAGTAAGCAACATGTAATGATGAGGCGGATCCACACAGAAAAAGGTAG 1140
Db      1081 GGGCTGGAAGTAAGCAACATGTAATGATGAGGCGGATCCACACAGAAAAAGGTAG 1140
QY      1141 ATCTGAATGCTGATCCCTGTGTGAGAGAAATAAGTAAGCAAGAAATCCATGCT 1200
Db      1141 ATCTGAATGCTGATCCCTGTGTGAGAGAAATAAGTAAGCAAGAAATCCATGCT 1200
QY      1201 CAGAGATTCCTAGAGATATGAGATGTTCTTGATTAACACTAAATACAGCATTCAGA 1260
Db      1201 CAGAGATTCCTAGAGATATGAGATGTTCTTGATTAACACTAAATACAGCATTCAGA 1260
QY      1261 AAGTTAATGATGTTTCCAGAGATGATGATGTTAGTTGTGATGACTCATATGATG 1320
Db      1261 AAGTTAATGATGTTTCCAGAGATGATGATGTTAGTTGTGATGACTCATATGATG 1320
QY      1321 GGGAGTCTGATCAAAATGCAAGAGTACATGATATTTGGACGTTTAAATGAGTGGATG 1380
Db      1321 GGGAGTCTGATCAAAATGCAAGAGTACATGATATTTGGACGTTTAAATGAGTGGATG 1380
QY      1381 AATATTCTGGTCTTACAGAGAAATAGACTTACTGGCCAGTATCTCTCATGAGGCTTTAA 1440
Db      1381 AATATTCTGGTCTTACAGAGAAATAGACTTACTGGCCAGTATCTCTCATGAGGCTTTAA 1440
```





QY 3601 AAGATAGTATTTGCTGAAAATGACATTAAAGAAAGTTCTGCTTTTATGACAAAGCG 3660  
DB 3601 AAGATAGTATTTGCTGAAAATGACATTAAAGAAAGTTCTGCTTTTATGACAAAGCG 3660  
QY 3661 TCCAGAAAGAGAGCTTAGCAGAGAGTCTAGCCCTTTCACCCATACACATTTGGCTCAGG 3720  
DB 3661 TCCAGAAAGAGAGCTTAGCAGAGAGTCTAGCCCTTTCACCCATACACATTTGGCTCAGG 3720  
QY 3721 GTTACGAGAGAGAGAGAGAAATTTAGAGTCTCTGAGAGAGAGTATCTAGTGGAGAG 3780  
DB 3721 GTTACGAGAGAGAGAGAGAGAAATTTAGAGTCTCTGAGAGAGAGTATCTAGTGGAGAG 3780  
QY 3781 AAGAGCTTCCCTGCTTCCACACCTTGTATTTGTTAAAGTAAACAAATATACCTTCTCAGT 3840  
DB 3781 AAGAGCTTCCCTGCTTCCACACCTTGTATTTGTTAAAGTAAACAAATATACCTTCTCAGT 3840  
QY 3841 CTACTAGGATAGCAGCCGTTGCTACCGAGTGTCTGTCTAAGAACACAGAGAGAAATTTAT 3900  
DB 3841 CTACTAGGATAGCAGCCGTTGCTACCGAGTGTCTGTCTAAGAACACAGAGAGAAATTTAT 3900  
QY 3901 TATCATGAGAAATAGCTTAATAGCTAGTAAACAGGATATATGCAAGGATGTC 3960  
DB 3901 TATCATGAGAAATAGCTTAATAGCTAGTAAACAGGATATATGCAAGGATGTC 3960  
QY 3961 AGGAACTCACTTAGTAGAGAAACAAATGTTGCTAGCTTGTCTTCTCAGAGTCA 4020  
DB 3961 AGGAACTCACTTAGTAGAGAAACAAATGTTGCTAGCTTGTCTTCTCAGAGTCA 4020  
QY 4021 GTGATTTGGAAGACTTGTACTGCTCAAAATACAAACCCAGAGATCTTCTTGATGGTCTT 4080  
DB 4021 GTGATTTGGAAGACTTGTACTGCTCAAAATACAAACCCAGAGATCTTCTTGATGGTCTT 4080  
QY 4081 CCAAAACAATGAGGATCAGTCTGAAAGCCAGGAGTGTGCTGAGTACAAAGAAATGG 4140  
DB 4081 CCAAAACAATGAGGATCAGTCTGAAAGCCAGGAGTGTGCTGAGTACAAAGAAATGG 4140  
QY 4141 TTTTCAGATGATGAAGAAAGAGAAAGGGCTTGAAGAAATATTCAGAGAGCAAAACA 4200  
DB 4141 TTTTCAGATGATGAAGAAAGAGAAAGGGCTTGAAGAAATATTCAGAGAGCAAAACA 4200  
QY 4201 TGGATTTAAACTTAGTGAAGACAGTCTGAGTGTGAGAGTGAAGAAAGCTCTGAG 4260  
DB 4201 TGGATTTAAACTTAGTGAAGACAGTCTGAGTGTGAGAGTGAAGAAAGCTCTGAG 4260  
QY 4261 ACTGCTAGAGGCTATCTCTCAGAGTGCATTTTAAACACTCAGAGAGGATACCATGC 4320  
DB 4261 ACTGCTAGAGGCTATCTCTCAGAGTGCATTTTAAACACTCAGAGAGGATACCATGC 4320  
QY 4321 AACATTAACCTGATTAAGCTCCAGAGAGAAATGGCTGAACCTAGAAAGCTGTGAACAGC 4380  
DB 4321 AACATTAACCTGATTAAGCTCCAGAGAGAAATGGCTGAACCTAGAAAGCTGTGAACAGC 4380  
QY 4381 ATGGAGAGGAGGCTTCTTAACAGCTACCTTCCATCATAGAGTACTCTTCTGAGG 4440  
DB 4381 ATGGAGAGGAGGCTTCTTAACAGCTACCTTCCATCATAGAGTACTCTTCTGAGG 4440  
QY 4441 ACCGTGCAAAATCCAGAAAGAGCAGATCAGAAAAAGAGTATTAACCTCAGAGAAAGTA 4500  
DB 4441 ACCGTGCAAAATCCAGAAAGAGCAGATCAGAAAAAGAGTATTAACCTCAGAGAAAGTA 4500  
QY 4501 GTGATTAACCTTAATAGCCAGAAATCCAGAGAGGCTTCTGTGACAAAGTTGAGGTGTG 4560  
DB 4501 GTGATTAACCTTAATAGCCAGAAATCCAGAGAGGCTTCTGTGACAAAGTTGAGGTGTG 4560  
QY 4561 CAGATTAATCTTACAGTAAAAATTAAGAACAGAGTGAAGAGTCACTCCCTTCTAAT 4620  
DB 4561 CAGATTAATCTTACAGTAAAAATTAAGAACAGAGTGAAGAGTCACTCCCTTCTAAT 4620  
QY 4621 GCCCATGATTAGATGATGAGGTGATGATGACAGTGTGCTGGAGTCTTCCAGAAATAGAA 4680  
DB 4621 GCCCATGATTAGATGATGAGGTGATGATGACAGTGTGCTGGAGTCTTCCAGAAATAGAA 4680

QY 4681 ACTACCATCTTCAAGAGAGGCTCATTAAGTTGTTGATGATGAGAGAGCAACAGCTGAG 4740  
DB 4681 ACTACCATCTTCAAGAGAGGCTCATTAAGTTGTTGATGATGAGAGAGCAACAGCTGAG 4740  
QY 4741 AGTGGGCGCACAGATTTTACGGAACATCTTACTTGGCCAGGCAAGATCTAGAGGAA 4800  
DB 4741 AGTGGGCGCACAGATTTTACGGAACATCTTACTTGGCCAGGCAAGATCTAGAGGAA 4800  
QY 4801 CCCCCTTACCTGGAATCTGGAATCAGCCCTTCTGAGTGAACCCGGAATCTGATCTG 4860  
DB 4801 CCCCCTTACCTGGAATCTGGAATCAGCCCTTCTGAGTGAACCCGGAATCTGATCTG 4860  
QY 4861 AAGACAGAGCCCAAGAGTCAAGTCTGTTGGCAACATACCAATCTTCAACCTCTGATTA 4920  
DB 4861 AAGACAGAGCCCAAGAGTCAAGTCTGTTGGCAACATACCAATCTTCAACCTCTGATTA 4920  
QY 4921 AAGTTCCCAATTTGAABTTGCAAGATCTGCCAGAGTCCAGCTGCTGCTCATTAATG 4980  
DB 4921 AAGTTCCCAATTTGAABTTGCAAGATCTGCCAGAGTCCAGCTGCTGCTCATTAATG 4980  
QY 4981 ATACTGCTGGGATTAATGCAATGGAAGAGTGTGAGAGAGAGAGCCAGAAATTCAG 5040  
DB 4981 ATACTGCTGGGATTAATGCAATGGAAGAGTGTGAGAGAGAGAGCCAGAAATTCAG 5040  
QY 5041 CTTCAACAGAAAGGCTCAACAAAGAAATGTCATGTTGCTGAGCTGAGCCAGAG 5100  
DB 5041 CTTCAACAGAAAGGCTCAACAAAGAAATGTCATGTTGCTGAGCTGAGCCAGAG 5100  
QY 5101 AATTATCTGCTGTGACAGTGTGCGAGAAACACACATCTTAACTTAATTA 5160  
DB 5101 AATTATCTGCTGTGACAGTGTGCGAGAAACACACATCTTAACTTAATTA 5160  
QY 5161 CTGAAGAGACTACTGATGTTGTAAGAAACAGATGCTGAGTGTGCTGTAACAGGAC 5220  
DB 5161 CTGAAGAGACTACTGATGTTGTAAGAAACAGATGCTGAGTGTGCTGTAACAGGAC 5220  
QY 5221 TGAATATTTTCTAGGAATTTGCGGAGAGAAATGGTAGTATTTCTGGGTGACC 5280  
DB 5221 TGAATATTTTCTAGGAATTTGCGGAGAGAAATGGTAGTATTTCTGGGTGACC 5280  
QY 5281 AGCTATTAAGAAAGAAATGCTGAATGAGTATGATTTTGAAGTCAAGAGAGATGTG 5340  
DB 5281 AGCTATTAAGAAAGAAATGCTGAATGAGTATGATTTTGAAGTCAAGAGAGATGTG 5340  
QY 5341 TCAATGGAAGAAACACCAAGGTCGCAAGAGAGAGAGATCCAGAGAGAAAGATCT 5400  
DB 5341 TCAATGGAAGAAACACCAAGGTCGCAAGAGAGAGAGATCCAGAGAGAAAGATCT 5400  
QY 5401 TCAGGGGCTAGAAATCTGTTGCTATGAGGCTTCCACCAATGCCCCAGATCACTGG 5460  
DB 5401 TCAGGGGCTAGAAATCTGTTGCTATGAGGCTTCCACCAATGCCCCAGATCACTGG 5460  
QY 5461 AATGAGTGTACACTGTGTGCTCTGTGAGAGAGCTTCTCATTTACCTG 5520  
DB 5461 AATGAGTGTACACTGTGTGCTCTGTGAGAGAGCTTCTCATTTACCTG 5520  
QY 5521 GCACAGGCTGTCACCAATTTGCTTGTGAGAGAGAGTGTGAGAGAGAGAGATGCT 5580  
DB 5521 GCACAGGCTGTCACCAATTTGCTTGTGAGAGAGAGTGTGAGAGAGAGAGATGCT 5580  
QY 5581 TCCATGCAATTTGSCAGATGTGTAGAGCAGCTGTGTGACCCGAGAGTGGTGTGACA 5640  
DB 5581 TCCATGCAATTTGSCAGATGTGTAGAGCAGCTGTGTGACCCGAGAGTGGTGTGACA 5640  
QY 5641 GTGATGACCTTACAGAGTGGAGAGAGTGTGACAGCTGATGATGATGATGATGATGAT 5700  
DB 5641 GTGATGACCTTACAGAGTGGAGAGAGTGTGACAGCTGATGATGATGATGATGATGAT 5700  
QY 5701 GCCACTACTGA 5711  
DB 5701 GCCACTACTGA 5711

Fri Jun 27 10:47:21 2003

Search completed: June 27, 2003, 01:01:50  
Job time : 14243.3 secs

us-09-734-672-3.rge

Page 45



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 27, 2003, 05:04:22 ; Search time 1769 Seconds  
(without alignments)  
1840.188 Million cell updates/sec

Title: US-09-734-672-3\_COPY\_2100\_2300  
Perfect score: 201  
Sequence: 1 AGGCACAGCAGAACTACA.....TTGTCATCTCAGCCTTCA 201

Scoring table: OLIGO\_NUC  
Gapop 60.0, Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size: 0  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :  
EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hlc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_plu:\*  
21: em\_gss\_vit:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	181	90.0	602 10 BE043993	BE043993 hK83c01.x
2	150	74.6	1089 13 BM452288	BM452288 AGENCOURT
3	48	23.9	443 14 B0308670	B0308670 MRO-B7600
4	44	21.2	14 025782	U25782 HSU25782 no
5	19	9.5	213 12 BG349328	BG349328 947028A03
6	19	9.5	219 9 A1414403	A1414403 ma399a01.x

7	19	9.5	387 12 BE367151	BE367151 RCI-GN002
8	19	9.5	412 10 AM531177	AM531177 UI-R-BR1-
9	19	9.5	444 12 BG349012	BG349012 947028A03
10	19	9.5	448 9 AA283425	AA283425 vb90e09.r
11	19	9.5	454 9 AA242194	AA242194 my30e01.r
12	19	9.5	493 14 B0031604	B0031604 UI-1-CFO-
13	19	9.5	514 9 AA178652	AA178652 mt18903.r
14	19	9.5	533 13 B1795365	B1795365 B1795365
15	19	9.5	563 13 BM045989	BM045989 603625142
16	19	9.5	580 9 A1504935	A1504935 v132c12.x
17	19	9.5	624 13 B1205435	B1205435 B1205435
18	19	9.5	2287 11 BC020362	BC020362 Mus muscu
19	18	9.0	259 12 BE560517	BE560517 UI-R-CO-h
20	18	9.0	293 14 T50377	T50377 yb55c05.s1
21	18	9.0	309 14 T51068	T51068 yb55c05.r1
22	18	9.0	376 17 CNS0288Y	AT232315 Tetradon
23	18	9.0	399 17 A0166402	A0166402 HS_3128_B
24	18	9.0	443 17 B54094	B54094 CIT-HSP-201
25	18	9.0	476 10 BE096207	BE096207 UI-R-BU0-
26	18	9.0	486 17 BH513308	BH513308 B0GJH12TR
27	18	9.0	512 14 B0130483	B0130483 i186h06.x
28	18	9.0	523 13 B1678186	B1678186 SMS86 SW
29	18	9.0	535 17 BH339312	BH339312 CH230-41D
30	18	9.0	540 17 A0947534	A0947534 Sheared D
31	18	9.0	544 9 AA964534	AA964534 UI-R-CO-g
32	18	9.0	554 17 TA291D05P	TA291D05P T. brucei
33	18	9.0	555 17 TA295E06P	TA295E06P T. brucei
34	18	9.0	640 9 A0077451	A0077451 A0077451
35	18	9.0	642 13 B1087604	B1087604 602852095
36	18	9.0	660 17 A2220818	A2220818 Sheared D
37	18	9.0	715 17 BH176278	BH176278 005.N.10-
38	18	9.0	715 17 CNS07J37	AT613237 T7 end of
39	18	9.0	723 17 AG181688	AG181688 Pan trogl
40	18	9.0	732 17 A2216722	A2216722 Sheared D
41	18	9.0	761 17 A0450104	A0450104 500008A05
42	18	9.0	837 17 A2199793	A2199793 SP_1040_B
43	18	9.0	862 17 A2677885	A2677885 EMT161TF
44	18	9.0	871 17 CNS059JH	AT327302 Tetradon
45	18	9.0	886 17 CNS0216T	AT196542 Tetradon

## ALIGNMENTS

RESULT 1  
BE043993/c 602 bp mRNA linear EST 08-JUN-2000  
LOCUS hK83c01.x1 NCI CGAP Lym12 Homo sapiens cDNA clone IMAGE:3003264 3'  
DEFINITION similar to gb:U14680 BREAST CANCER TYPE 1 SUSCEPTIBILITY PROTEIN (HUMAN); mRNA sequence.

ACCESSION BE043993  
VERSION BE043993.1 GI:8361046  
KEYWORDS EST.

## ORGANISM

human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

## REFERENCE

1 (bases 1 to 602)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index

## TITLE

## JOURNAL

## COMMENT

Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapb-remail.nih.gov  
Life Technologies catalog #: 11547-015  
DNA sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL, send email to:  
info@image.lnl.gov

Seq primer: -400p from Glibco  
High quality sequence stop: 418.  
Location/Qualifiers

FEATURES  
SOURCE:  
1. 602

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3003264"
/clone_lib="NCI CGAP Lym12"
/tissue_type="lymphoma, follicular mixed small and large
cell"
/lab_host="DH10B"
/Note="Organ: lymph node; Vector: pCMV-SPORT6; Site:1:
SalI; Site:2: NotI; Cloned unidirectionally. Primer:
Oligo dT. Average insert size 1.25 kb. Life Technologies
catalog #: 11547-015"

BASE COUNT      131 a      114 c      113 g      243 t      1 others
ORIGIN
Query Match      90.0%; Score 181; DB 10; Length 602;
Best Local Similarity 100.0%; Pred. No. 3.4e-84;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 21 ACTCATGGAAGTAAAGAACTGCAACTGAGCCAGAGAGTAAAGCAATGACAA 80
    |||||||
DB 188 ACTCATGGAAGTAAAGAACTGCAACTGAGCCAGAGAGTAAAGCAATGACAA 129
    |||||||
OY 81 GACAGTAAAGACATGATGATCTTCCAGAGCTGAATTAACAATGACCTGG 140
    |||||||
DB 128 GACAGTAAAGACATGATGATCTTCCAGAGCTGAATTAACAATGACCTGG 69
    |||||||
OY 141 TTCTTTTAAGTGTTCATATCCAGTAACTTAAGATTTGTCAATCTGACCTTC 200
    |||||||
DB 68 TTCTTTTAAGTGTTCATATCCAGTAACTTAAGATTTGTCAATCTGACCTTC 9
    |||||||
OY 201 A 201
    |
DB 8 A 8

RESULT 2
BM452288      1089 bp      mRNA      linear      EST 05-FEB-2002
LOCUS      AGENCOURT_6386302 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5526666
DEFINITION      5', mRNA sequence.
ACCESSION      BM452288
VERSION      BM452288.1 GI:16501328
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1. (bases 1 to 1089)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCRD/FTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LLM12200 row: k column: 19
High quality sequence stop: 667.
Location/Qualifiers
1. 1089
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5526666"
/clone_lib="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/Note="Organ: skin; Vector: pCMV-SPORT6; Site:1: NotI;
Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life

```

```

Technologies."
BASE COUNT      389 a      212 c      248 g      240 t
ORIGIN
Query Match      74.6%; Score 150; DB 13; Length 1089;
Best Local Similarity 99.5%; Pred. No. 7.5e-68;
Matches 200; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 AGGCACAGCAGAACTTACATGATGAGGTAAGAACTGCACTGAGCCAAAG 60
    |||||||
DB 188 AGGCACAGCAGAACTTACATGATGAGGTAAGAACTGCACTGAGCCAAAG 247
    |||||||
OY 61 AGTAACAGCAGAACTTACATGATGAGGTAAGAACTGCACTGAGCCAAAG 120
    |||||||
DB 248 AGTAACAGCAGAACTTACATGATGAGGTAAGAACTGCACTGAGCCAAAG 307
    |||||||
OY 121 AAGTTAAACAATGACCTGCTTTTACTTAAGTGTTCATTAACAGTAATTAAGA 180
    |||||||
DB 308 AAGTTAAACAATGACCTGCTTTTACTTAAGTGTTCATTAACAGTAATTAAGA 367
    |||||||
OY 181 TTGTCAATGCTGACCTTCCA 201
    |||||||
DB 368 TTGTCAATGCTGACCTTCCA 388

RESULT 3
B0308670      443 bp      mRNA      linear      EST 16-MAY-2002
LOCUS      MR0-BT6000-020701-101-b03 BT6000 Homo sapiens cDNA, mRNA sequence.
DEFINITION      B0308670
ACCESSION      B0308670.1 GI:20850612
VERSION      B0308670.1
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1. (bases 1 to 443)
Dias Neto,E., Garcia Correia,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bais,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=MR0&t2=MR0-BT6000-
020701-101-b03&t3=2001-07-02&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 355.
Location/Qualifiers
1. 443
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="BT6000"
/dev_stage="Adult"
/Note="Organ: breast; Vector: puc18; Site:1: SmaI; Site:2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
,716 - Ludwig Institute for Cancer Research) profiles
into the puc 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low

```





High quality sequence stop: 206.

# FEATURES

source

Location/Qualifiers

1. 219  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="IMAGE:318792"  
/dev\_stage="19.5 dpc total fetus"  
/lab\_host="DH10B (ampicillin resistant)"  
/note="Vector: pT73D (Pharmacia) with a modified polylinker. Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAGTGGAGCGCCGCAATTTTCTTTTCTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73D vector (Pharmacia). Library went through one round of normalization to a cot = 5. Library constructed by Bento Soares and M. Fatima Boudado. RNA was kindly provided by Dr. Minoru Ko (Wayne State University)."

## BASE COUNT

61 a 46 c 32 g 80 t

## ORIGIN

Query Match 9.5%; Score 19; DB 9; Length 219;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

143 CTTTACTAGTGTCAAA 161

21 CTTTACTAGTGTCAAA 3

Db

RESULT 7  
BF367151 387 bp mRNA linear EST 24-NOV-2000  
LOCUS RCI-GN0021-300800-022-cl1 GN0021 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BF367151  
VERSION BF367151.1 GI:11329176  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 387)  
Dias Neto, E., Garcia Correia, R., Verjovski-Almeida, S., Briones, M.R.,  
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
Goldman, G.H., Carvalho, A.F., Matsushima, A., Bala, G.S., Simpson, D.H.,  
Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,  
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
Simpson, A.J.

Shogun sequencing of the human transcriptome with ORF expressed  
sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20020663  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=RC1&c2=RC1-GN0021-  
300800-022-cl1&c3=2000-08-30&t4=1)  
Seq primer: puc 18 forward  
High quality sequence stop: 387.

## FEATURES

source

Location/Qualifiers  
1. 387  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="GN0021"

/dev\_stage="Adult"

/note="Organ: placenta-normal; Vector: puc18; Site\_1: SmaI  
; Site\_2: SmaI; A mini-library was made by cloning  
products derived from ONESIES PCR (U.S. Letters Patent  
Application No. 196,716 - Ludwig Institute for Cancer  
Research) profiles into the puc 18 vector. Reverse  
transcription of tissue mRNA and cDNA amplification were  
performed under low stringency conditions."

## BASE COUNT

90 a 108 c 52 g 137 t

## ORIGIN

Query Match 9.5%; Score 19; DB 12; Length 387;  
Best Local Similarity 100.0%; Pred. No. 37;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

128 CAATGCACTGCTCTTT 146

186 CAATGCACTGCTCTTT 204

Db

RESULT 8  
AM531177/c 412 bp mRNA linear EST 06-MAR-2000  
LOCUS UI-R-BT1-akv-e-12-0-UI-s1 UI-R-BT1 Rattus norvegicus cDNA clone  
DEFINITION UI-R-BT1-akv-e-12-0-UI 3', mRNA sequence.  
ACCESSION AM531177  
VERSION AM531177.1 GI:7173591  
KEYWORDS EST.  
SOURCE Norway rat.  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE 1 (bases 1 to 412)  
Boudado, M.F., Lennon, G. and Soares, M.B.  
Normalization and subtraction: two approaches to facilitate gene  
discovery  
Genome Res. 6 (9), 791-806 (1996)  
97044477  
Contact: Soares, MB  
Program for Rat Gene Discovery and Mapping  
University of Iowa  
451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: mssoares@iuiowa.edu

The sequence contained an oligo-dT track that was present in the  
oligonucleotide that was used to prime the synthesis of first  
strand cDNA and therefore this may represent a bonafide poly A  
tail. The sequence tag present in the cDNA between the NotI site  
and the oligo-dT track served to identify it as a clone from the  
normalized thalamus library cDNA library Preparation: M.B. Soares  
Lab Clone distribution: clones will be available through Research  
Genetics (www.resgen.com)  
Seq primer: M13 Forward  
POLYA=Yes.

## FEATURES

source

Location/Qualifiers  
1. 412  
/organism="Rattus norvegicus"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
/clone="UI-R-BT1-akv-e-12-0-UI"  
/clone\_lib="UI-R-BT1"  
/dev\_stage="adult"  
/lab\_host="DH10B (Life Technologies)"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker. Site\_1: Not I; Site\_2: Eco RI; The library  
UI-R-BT1 is a subtracted library derived from a mixture of  
the following tissues: hippocampus, thalamus, mid-brain,  
medulla, corpus striatum, cerebral cortex and testis. For  
a detailed description of the library from which this  
clone was derived, please visit our web site at  
ratest.eng.uiowa.edu. The subtraction has been previously

described in (Bonaldo, Lennon and Soares; Genome Research  
6:791-806, 1996)  
TAG\_LIB=UI-R-BT1  
TAG\_TISSUE=thalamus  
TAG\_SEQ=GATCG\*

BASE COUNT 109 a 68 c 72 g 163 t  
ORIGIN

Query Match 9.5%; Score 19; DB 10; Length 412;  
Best Local Similarity 100.0%; Pred. No. 38;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 142 TCTTTACTAGTGTCAA 160  
|||||  
Db 404 TCTTTACTAGTGTCAA 386

RESULT 9  
BG349012/c 444 bp mRNA linear EST 01-MAR-2001  
LOCUS 947028A03.x2 947 - 2 week shoot from Barkan lab Zea mays cDNA, mRNA  
DEFINITION  
SEQUENCE  
ACCESSION BG349012 GI:13177662  
VERSION BG349012.1  
KEYWORDS EST.  
SOURCE Zea mays.  
ORGANISM Zea mays.

REFERENCE  
AUTHORS Walbot,V.  
TITLE Mize ESTs from various cDNA libraries sequenced at Stanford University  
JOURNAL Unpublished (1999)  
COMMENT Contact: Walbot V  
Department of Biological Sciences  
Stanford University  
855 California Ave, Palo Alto, CA 94304, USA  
Tel: 650 723 2227  
Fax: 650 725 8221  
Email: walbot@stanford.edu  
Plate: 947028 row: A column: 03.  
Location/Qualifiers

FEATURES  
source  
1..444  
/organism="Zea mays"  
/cultivar="B73"  
/db\_xref="taxon:4577"  
/clone\_lib="947 - 2 week shoot from Barkan lab"  
/tissue\_type="leaf and stem, including leaf base"  
/dev\_stage="2 week old seedling (3 leaves)"  
/lab\_host="X11-Blue"  
/note="Organ: Shoot; Vector: Lambda ZAP (pBluescript SK-);  
Site\_1: EcoRI; Site\_2: XhoI; Directionally cloned using  
stratagene's Unizap XR cDNA cloning kit with the 5' end  
at the EcoRI site. The library represents 8 x 10e5  
independent recombinant phage. The plants were greenhouse  
grown."

BASE COUNT 148 a 116 c 75 g 105 t  
ORIGIN

Query Match 9.5%; Score 19; DB 12; Length 444;  
Best Local Similarity 100.0%; Pred. No. 38;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 64 AACAGCAATGACAGA 82  
|||||  
Db 436 AACAGCAATGACAGA 418

RESULT 10  
AA285425 448 bp mRNA linear EST 08-APR-1997  
LOCUS

DEFINITION vb90e09.r1 Soares mouse 3nbms Mus musculus cDNA clone IMAGE:764296  
5', mRNA sequence.  
ACCESSION AA285425  
VERSION AA285425.1 GI:1929816  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus

REFERENCE  
AUTHORS Mammalia: Eutheria; Rodentia; Sciurognath; Muridae; Murinae; Mus.  
1 (bases 1 to 448)  
Marr, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,  
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,  
Schellenberg, K., Stepien, M., Tan, F., Underwood, K., Moore, B.,  
Theising, B., Wyllie, T., Lennon, G., Soares, B., Wilson, R. and  
Waterston, R.

THE WASHU-HHMI MOUSE EST PROJECT  
Unpublished (1996)  
Contact: Maria M/Mouse EST Project  
Washu-HHMI Mouse EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@wustl.edu

This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.lnl.gov) for further information.  
MGI:465216

Seq primer: -28ml3 rev2 ET from Amersham  
High quality sequence stop: 432.

FEATURES  
source  
1..448  
Location/Qualifiers

/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone\_lib="IMAGE:764296"  
/clone\_lib="Soares mouse 3nbms"  
/sex="male"  
/tissue\_type="Spleen"  
/dev\_stage="4 weeks"  
/lab\_host="DH10B"  
/note="Vector: p773D-Pac (Pharmacia) with a modified  
polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA  
was primed with a Not I - oligo(dT) primer [5'  
TGTTCACCAATCGACGAGCGCGCGCTTTTCTTTTCTTTTCTTTTCTTTT  
3']; double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified p773 vector. RNA  
provided by Dr. Bertrand Jordan. Library went through  
three rounds of normalization, and was constructed by  
Bento Soares and M.Fatima Bonaldo."

BASE COUNT 165 a 88 c 75 g 120 t  
ORIGIN

Query Match 9.5%; Score 19; DB 9; Length 448;  
Best Local Similarity 100.0%; Pred. No. 38;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 142 TCTTTACTAGTGTCAA 160  
|||||  
Db 88 TCTTTACTAGTGTCAA 106

RESULT 11  
AA242194 454 bp mRNA linear EST 07-MAR-1997  
LOCUS my30e01.r1 Barstead mouse pooled organs MRLR4 Mus musculus cDNA  
DEFINITION clone IMAGE:697368 5', mRNA sequence.  
ACCESSION AA242194  
VERSION AA242194.1 GI:1873385  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

FEATURES	Location/Qualifiers									
SOURCE	1.493	/organism="Mus musculus"								
	/strain="C57BL/6J"									
	/db_xref="taxon:10090"									
	/clone="UI-1-CF0-ang-c-08-0-UI"									
	/clone.lib="NCI CGAP.P1R1"									
	/tissue.type="Trophoblast"									
	/lab_host="DH10B (Life Technologies)"									
	/note="Organ: Placenta; Vector: pT7f3-Pac (Pharmacia) with a modified polylinker; Site_1: Ecor I; Site_2: Not I; NCI CGAP.P1R1 is a subcloned cDNA library constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an Ecor I adaptor, digested with Not I, and cloned directionally into pT7f3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tags for this library are GTGGG, ATCAT, GGGTG, GGTTC. For additional information, contact: Bento Soares, bento-soares@uiowa.edu									
BASE COUNT	116 a	123 c	99 g	155 t	5 days					
ORIGIN	TAG_TISSUE=Placenta mouse 11.5-12.5 days									
	TAG_SEQ=ATCAT"									
Query Match	Best Local Similarity 9.5%; Score 19; DB 14; Length 493;									
	Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
OY	143	CTTTTACTAGTGTCANA 161								
Db	34	CTTTTACTAGTGTCANA 16								
RESULT 13										
LOCUS	AA178652	514 bp	MRNA	linear	EST 17-FEB-1997					
DEFINITION	mtl1g03.r1 Soares mouse 3mbms Mus musculus cDNA clone IMAGE:621460									
ACCESSION	AA178652									
VERSION	AA178652.1	GI:1760020								
KEYWORDS	EST.									
SOURCE	house mouse.									
ORGANISM	Mus musculus									
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciuromorphi; Muridae; Murinae; Mus. 1 (bases 1 to 514)									
AUTHORS	Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kuba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.									
TITLE	The WashU-HHMI Mouse EST Project									
JOURNAL	Unpublished (1996)									
COMMENT	Contact: Marra M/Mouse EST Project WashU-HHMI Mouse EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@watson.wustl.edu This clone is available royalty-free through LNL; contact the									

IMAGE Consortium (infoimage.llnl.gov) for further information.  
 MG1:382284  
 Putative full length read  
 vector to vector length is 635  
 Seq primer: -28M13 rev2 from Amersham  
 High quality sequence stop: 480.  
 Location/Qualifiers

FEATURES  
 SOURCE

1. 514  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:621460"  
 /clone\_lib="Soares mouse 3BDM5"  
 /sex="male"  
 /tissue\_type="Spleen"  
 /dev\_stage="4 weeks"  
 /lab\_host="DH10B"  
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified  
 polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA  
 was primed with a Not I - oligo(dt) primer [5',  
 TGTTCACATCTGAGGAGGAGCGCGCGTGTGTGTGTGTGTGTGTGT  
 3']; double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not I  
 and Eco RI sites of the modified pT7T3 vector. RNA  
 provided by Dr. Bertrand Jordan. Library went through  
 three rounds of normalization, and was constructed by  
 Bento Soares and M.Fatima Bonaldo."  
 BASE COUNT 186 a 98 c 90 g 140 t  
 ORIGIN

Query Match 9.5%; Score 19; DB 9; Length 514;  
 Best Local Similarity 100.0%; Pred. No. 40;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 142 TCTTTTACTAGTGTCA 160  
 ||||||||||||||||  
 Db 161 TCTTTTACTAGTGTCA 179

RESULT 14  
 BUI95365/c 533 bp mRNA linear EST 24-JAN-2002  
 LOCUS BUI95365 normalized full length cDNA library, chloronemata,  
 DEFINITION caulonemata and rhizoid-like protonemata Physcomitrella patens  
 subsp. patens cDNA clone pphn23d17 5', mRNA sequence.  
 BUI95365  
 BUI95365.1 GI:18363293

EST.  
 Physcomitrella patens subsp. patens.  
 Physcomitrella patens subsp. patens  
 Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Bryophyta;  
 Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella;  
 1 (bases 1 to 533)  
 Fujita,T., Shin-I,T., Seki,M., Kamiya,A., Uchiyama,I., Nishiyama,T.,  
 Carninci,P., Hayashizaki,Y., Shinozaki,K., Kohara,Y. and Hasebe  
 /M.

TITLE  
 Comparison of the moss Physcomitrella patens genome with flowering  
 plants genome  
 JOURNAL  
 Unpublished (2002)  
 COMMENT  
 Contact: Tadasu Shin-I  
 Center for Genetic Resource Information  
 National Institute of Genetics  
 111 Yata, Mishima, Shizuoka 411-8540, Japan  
 Tel: 81-559-81-6856  
 Fax: 81-559-81-6855

EMAIL: tshin@genes.nig.ac.jp  
 A backbone of the vector is pBluescript II, that was in vivo  
 excised from a modified lps phage vector (Mo bi Tec, Germany). XhoI  
 digested-5' end of cDNA is ligated to SalI site of the vector, and  
 the BamHI digested-3' end including poly-A tail is ligated to BamHI  
 site of the vector. cDNA insert could be amplified with  
 conventional T7 and T3 primers. This normalized full-length cDNA  
 library was generated basically according to the method described

in Genome Research 10, 1617-1630 (2000), Carninci, P. et al.  
 Protonemata were blended by the POLYTRON, and then cultivated on  
 the BCD medium containing 1µM NAA (naphthalene acetic acid) for 8  
 to 11 days under the continuous light.  
 Location/Qualifiers

FEATURES  
 SOURCE

1. 533  
 /organism="Physcomitrella patens subsp. patens"  
 /db\_xref="taxon:145481"  
 /clone="pphn23d17"  
 /clone\_lib="normalized full length cDNA library,  
 chloronemata, caulonemata and rhizoid-like protonemata"  
 /tissue\_type="mixture of chloronemata, caulonemata and  
 rhizoid-like protonemata"  
 BASE COUNT 121 a 110 c 162 g 140 t  
 ORIGIN

Query Match 9.5%; Score 19; DB 13; Length 533;  
 Best Local Similarity 100.0%; Pred. No. 40;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 53 CCAGAGAGGACGACGCC 71  
 ||||||||||||||||  
 Db 239 CCAGAGAGGACGACGCC 221

RESULT 15  
 BMO45989 563 bp mRNA linear EST 07-NOV-2001  
 LOCUS BMO45989  
 DEFINITION 603625142F1 NIH\_MGC\_40 Homo sapiens cDNA clone IMAGE:5451611 5',  
 mRNA sequence.  
 BMO45989  
 BMO45989.1 GI:16775256  
 EST.  
 SOURCE  
 HUMAN.  
 Human.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 563)  
 NIH-MGC http://mgc.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs@mail.nih.gov  
 Tissue Procurement: DCTD/DTP  
 CDNA Library Preparation: Ling Hong/Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LMNL at:  
 http://image.llnl.gov  
 Plate: LHCN1942 row: d column: 12  
 High quality sequence stop: 389.  
 Location/Qualifiers

1. 563  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5451611"  
 /clone\_lib="NIH\_MGC\_40"  
 /tissue\_type="Carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: prostate; Vector: pOTB7; Site\_1: XhoI;  
 Site\_2: EcoRI. cDNA made by oligo-dT priming.  
 Directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGCACGAG(G). Library constructed by  
 Ling Hong in the Laboratory of Gerald M. Rubin (University  
 of California, Berkeley) using Zap-CDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH\_MGC Library."

TITLE  
 Comparison of the moss Physcomitrella patens genome with flowering  
 plants genome  
 JOURNAL  
 Unpublished (2002)  
 COMMENT  
 Contact: Tadasu Shin-I  
 Center for Genetic Resource Information  
 National Institute of Genetics  
 111 Yata, Mishima, Shizuoka 411-8540, Japan  
 Tel: 81-559-81-6856  
 Fax: 81-559-81-6855

EMAIL: tshin@genes.nig.ac.jp  
 A backbone of the vector is pBluescript II, that was in vivo  
 excised from a modified lps phage vector (Mo bi Tec, Germany). XhoI  
 digested-5' end of cDNA is ligated to SalI site of the vector, and  
 the BamHI digested-3' end including poly-A tail is ligated to BamHI  
 site of the vector. cDNA insert could be amplified with  
 conventional T7 and T3 primers. This normalized full-length cDNA  
 library was generated basically according to the method described

Matches	19;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	75	TGACACAGACAGTAAAGA	93						
Db	410	TGACACAGACAGTAAAGA	428						

Search completed: June 27, 2003, 06:35:34  
 Job time : 1787 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 26, 2003, 17:13:18 ; Search time 1905 Seconds

(without alignments)  
3070.690 Million cell updates/sec

Title: US-09-734-672-3\_COPY\_2100\_2300

Perfect score: 201

Sequence: 1 AGGCACAGCAGAACTACA.....TTGTCAATCTAGCCTTCA 201

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapept 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenBank:\*

1: gb\_ba:\*

2: gb\_bg:\*

3: gb\_in:\*

4: gb\_cm:\*

5: gb\_ov:\*

6: gb\_pat:\*

7: gb\_ph:\*

8: gb\_pl:\*

9: gb\_pr:\*

10: gb\_ro:\*

11: gb\_sts:\*

12: gb\_sy:\*

13: gb\_un:\*

14: gb\_vl:\*

15: em\_ba:\*

16: em\_fun:\*

17: em\_hum:\*

18: em\_in:\*

19: em\_mu:\*

20: em\_om:\*

21: em\_or:\*

22: em\_ov:\*

23: em\_pat:\*

24: em\_ph:\*

25: em\_pl:\*

26: em\_ro:\*

27: em\_sts:\*

28: em\_un:\*

29: em\_vl:\*

30: em\_hlg\_hum:\*

31: em\_hlg\_inv:\*

32: em\_hlg\_other:\*

33: em\_hlg\_mus:\*

34: em\_hlg\_pln:\*

35: em\_hlg\_rtd:\*

36: em\_hlg\_mam:\*

37: em\_hlg\_vrl:\*

38: em\_sy:\*

39: em\_hlgo\_hum:\*

40: em\_hlgo\_mus:\*

41: em\_hlgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	201	100.0	5711	6 AR007333	AR007333 Sequence
2	201	100.0	5711	6 AR007335	AR007335 Sequence
3	201	100.0	5711	6 AR112808	AR112808 Sequence
4	201	100.0	5711	6 AR112809	AR112809 Sequence
5	201	100.0	5711	6 AR112809	AR112809 Sequence
6	150	74.6	298	9 AF110102	AF110102 Homo sapi
7	150	74.6	649	9 AF274503	AF274503 Homo sapi
8	150	74.6	4249	6 AR004691	AR004691 Sequence
9	150	74.6	4249	6 AR008177	AR008177 Sequence
10	150	74.6	4249	6 AR136960	AR136960 Sequence
11	150	74.6	4249	6 AR176961	AR176961 Sequence
12	150	74.6	4249	6 AR0956	AR0956 Sequence
13	150	74.6	4249	6 AR181052	AR181052 Sequence
14	150	74.6	5656	6 AR048657	AR048657 Sequence
15	150	74.6	5656	6 AR048657	AR048657 Sequence
16	150	74.6	5689	6 AR048659	AR048659 Sequence
17	150	74.6	5689	6 AR048659	AR048659 Sequence
18	150	74.6	5693	9 AF005068	AF005068 Homo sapi
19	150	74.6	5707	6 AR048667	AR048667 Sequence
20	150	74.6	5707	6 AR048658	AR048658 Sequence
21	150	74.6	5709	6 AR048663	AR048663 Sequence
22	150	74.6	5709	6 AR048663	AR048663 Sequence
23	150	74.6	5709	6 AR048664	AR048664 Sequence
24	150	74.6	5709	6 AR048665	AR048665 Sequence
25	150	74.6	5709	6 AR281859	AR281859 Sequence
26	150	74.6	5709	6 AR0793	AR0793 Sequence
27	150	74.6	5709	6 AR0798	AR0798 Sequence
28	150	74.6	5709	6 AR0799	AR0799 Sequence
29	150	74.6	5709	6 AR0800	AR0800 Sequence
30	150	74.6	5710	6 AR048662	AR048662 Sequence
31	150	74.6	5710	6 AR0797	AR0797 Sequence
32	150	74.6	5711	6 AR007334	AR007334 Sequence
33	150	74.6	5711	6 AR033056	AR033056 Sequence
34	150	74.6	5711	6 AR048660	AR048660 Sequence
35	150	74.6	5711	6 AR048666	AR048666 Sequence
36	150	74.6	5711	6 AR112810	AR112810 Sequence
37	150	74.6	5711	6 AR0795	AR0795 Sequence
38	150	74.6	5711	6 AR0801	AR0801 Sequence
39	150	74.6	5711	9 HSU14680	U14680 Homo sapien
40	150	74.6	5712	6 AR048668	AR048668 Sequence
41	150	74.6	5712	6 AR070223	AR070223 Sequence
42	150	74.6	5712	6 AR118507	AR118507 Sequence
43	150	74.6	5712	6 AR125601	AR125601 Sequence
44	150	74.6	5712	6 AR184044	AR184044 Sequence
45	150	74.6	5712	6 AR0803	AR0803 Sequence

#### ALIGNMENTS

RESULT 1

LOCUS AR007333

DEFINITION Sequence 1 from patent US 5750400.

ACCESSION AR007333

VERSION AR007333.1

KEYWORDS GI:3966817

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 5711)

AUTHORS Murphy,P.D., Allen,A.C., Alvares,C.P., Crltz,B.S., Olson,S.J., Scheitler,D.B. and Zeng,B.

TITLE Coding sequences of the human BRCA1 gene

JOURNAL Patent: US 5750400-A 1 12-MAY-1998;

FEATURES	Location/Qualifiers
source	1..5711
BASE COUNT	1953 a 1099 c 1277 g 1382 t
ORIGIN	
Query Match	100.0%; Score 201; DB 6; Length 5711;
Best Local Similarity	100.0%; Pred. No. 5.3e-103; Mismatches 0; Indels 0; Gaps 0;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 AGGCACAGCAAAACCTACAACTCATGGAAGGTAAGAAGCACTGCACCTGAGCCAAAG 60
DB	2100 AGGCACAGCAAAACCTACAACTCATGGAAGGTAAGAAGCACTGCACCTGAGCCAAAG 2159
QY	61 AGTAACAAGCCAAATGAAACAGACAGCAAGTAAAGACATGACAGTACTTTCCAGAGCTG 120
DB	2160 AGTAACAAGCCAAATGAAACAGACAGCAAGTAAAGACATGACAGTACTTTCCAGAGCTG 2219
QY	121 AAGTTAACAAATGACCTGTTCTTTACTAAGTGTTCAAATACACAGTGAACCTTAAGAA 180
DB	2220 AAGTTAACAAATGACCTGTTCTTTACTAAGTGTTCAAATACACAGTGAACCTTAAGAA 2279
QY	181 TTTGTCAATCCTAGCCTTCCA 201
DB	2280 TTTGTCAATCCTAGCCTTCCA 2300
RESULT 2	
AR007335	5711 bp DNA linear PAT 04-DEC-1998
LOCUS	AR007335
DEFINITION	Sequence 5 from patent US 5750400.
ACCESSION	AR007335
VERSION	AR007335.1 GI:3966819
KEYWORDS	
SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	1 (bases 1 to 5711)
AUTHORS	Murphy,P.D., Allen,A.C., Alvares,C.P., Critz,B.S., Olson,S.J., Schelker,D.B. and Zeng,B.
TITLE	Coding sequences of the human BRCA1 gene
JOURNAL	Patent: US 5750400-A 5 12-MAY-1998;
FEATURES	Location/Qualifiers
source	1..5711
BASE COUNT	1953 a 1098 c 1277 g 1383 t
ORIGIN	
Query Match	100.0%; Score 201; DB 6; Length 5711;
Best Local Similarity	100.0%; Pred. No. 5.3e-103; Mismatches 0; Indels 0; Gaps 0;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 AGGCACAGCAAAACCTACAACTCATGGAAGGTAAGAAGCACTGCACCTGAGCCAAAG 60
DB	2100 AGGCACAGCAAAACCTACAACTCATGGAAGGTAAGAAGCACTGCACCTGAGCCAAAG 2159
QY	61 AGTAACAAGCCAAATGAAACAGACAGCAAGTAAAGACATGACAGTACTTTCCAGAGCTG 120
DB	2160 AGTAACAAGCCAAATGAAACAGACAGCAAGTAAAGACATGACAGTACTTTCCAGAGCTG 2219
QY	121 AAGTTAACAAATGACCTGTTCTTTACTAAGTGTTCAAATACACAGTGAACCTTAAGAA 180
DB	2220 AAGTTAACAAATGACCTGTTCTTTACTAAGTGTTCAAATACACAGTGAACCTTAAGAA 2279
QY	181 TTTGTCAATCCTAGCCTTCCA 201
DB	2280 TTTGTCAATCCTAGCCTTCCA 2300
RESULT 3	
AR112808	5711 bp DNA linear PAT 16-MAY-2001
LOCUS	AR112808
DEFINITION	Sequence 1 from patent US 6130322.

ACCESSION	ARL12808
VERSION	ARL12808.1 GI:14092708
KEYWORDS	
SOURCE	Unknown.
ORGANISM	Unclassified.
REFERENCE	1 (bases 1 to 5711)
AUTHORS	Murphy,P.D., Allen,A.C.P., Alvares,C.P., Critz,B.S., Olson,S.J., Thumber,D. and Zeng,B. Coding sequences of the human BRCA1 gene Patent: US 6130322-A 1 10-OCT-2000; Location/Qualifiers 1..5711
JOURNAL	
FEATURES	source /organism="unknown"
BASE COUNT	1953 a 1099 c 1277 g 1382 t
ORIGIN	
Query Match	100.0%; Score 201; DB 6; Length 5711;
Best Local Similarity	100.0%; Pred. No. 5.3e-103;
Matches 201:	Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	1 AGGCACAGCAGAAGAACTCATCACTATGCAAGGTAAAGAACCTGCAACTGGAGCCAAAGAAG 60
Dd	
Db	2100 AGGCACAGCAGAAGAACTCATCACTATGCAAGGTAAAGAACCTGCAACTGGAGCCAAAGAAG 2159
OY	61 AGTAAACAAGCCAAATGAACAGACAGCAAGTAATAAAGACATACAGTAGTACTTTCCAGAGCTG 120
Dd	
Db	2160 AGTAAACAAGCCAAATGAACAGACAGCAAGTAATAAAGACATACAGTAGTACTTTCCAGAGCTG 2219
OY	121 AAGTTAAACAATGACACCTGGTGTCCTTTTACTAAGTGTTCAAATFACCAGTGACTTAAGAA 180
Dd	
Db	2220 AAGTTAAACAATGACACCTGGTGTCCTTTTACTAAGTGTTCAAATFACCAGTGACTTAAGAA 2279
OY	181 TTTGTCAATCGTAGCCCTTCCA 201
Dd	
Db	2280 TTTGTCAATCGTAGCCCTTCCA 2300
OY	
RESULT 4	
LOCUS	ARL12809 5711 bp DNA linear PAT 16-MAY-2001
DEFINITION	Sequence 3 from patent US 6130322.
ACCESSION	ARL12809
VERSION	ARL12809.1 GI:14092709
KEYWORDS	
SOURCE	Unknown.
ORGANISM	Unclassified.
REFERENCE	1 (bases 1 to 5711)
AUTHORS	Murphy,P.D., Allen,A.C.P., Alvares,C.P., Critz,B.S., Olson,S.J., Thumber,D. and Zeng,B. Coding sequences of the human BRCA1 gene Patent: US 6130322-A 3 10-OCT-2000; Location/Qualifiers 1..5711
JOURNAL	
FEATURES	source /organism="unknown"
BASE COUNT	1953 a 1098 c 1277 g 1383 t
ORIGIN	
Query Match	100.0%; Score 201; DB 6; Length 5711;
Best Local Similarity	100.0%; Pred. No. 5.3e-103;
Matches 201:	Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	1 AGGCACAGCAGAAGAACTCATCACTATGCAAGGTAAAGAACCTGCAACTGGAGCCAAAGAAG 60
Dd	
Db	2100 AGGCACAGCAGAAGAACTCATCACTATGCAAGGTAAAGAACCTGCAACTGGAGCCAAAGAAG 2159
OY	61 AGTAAACAAGCCAAATGAACAGACAGCAAGTAATAAAGACATACAGTAGTACTTTCCAGAGCTG 120
Dd	
Db	2160 AGTAAACAAGCCAAATGAACAGACAGCAAGTAATAAAGACATACAGTAGTACTTTCCAGAGCTG 2219
OY	121 AAGTTAAACAATGACACCTGGTGTCCTTTTACTAAGTGTTCAAATFACCAGTGACTTAAGAA 180
Dd	
Db	2220 AAGTTAAACAATGACACCTGGTGTCCTTTTACTAAGTGTTCAAATFACCAGTGACTTAAGAA 2279



Db 2220 AAGTTACCAATGCACTGGTCTTTTACTAAGTTCATAATACAGTGAATTAAAGCA 2279  
Oy 181 TTGTCAATCCTAGCCTTCCA 201  
|||||  
Db 2280 TTGTCAATCCTAGCCTTCCA 2300

RESULT 5  
LOCUS 159546 5711 bp DNA linear PAT 07-OCT-1997  
DEFINITION Sequence 1 from patent US 5654155.  
ACCESSION 159546  
VERSION 159546.1 GI:2478178  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 5711)  
AUTHORS Murphy,P.D., Allen,A.C., Alvares,C.P., Critz,B.S., Olson,S.J.,  
Schelter,D.B. and Zeng,B.  
TITLE Consensus sequence of the human BRCA1 gene  
JOURNAL Patent: US 5654155-A 1 05-AUG-1997;  
FEATURES  
Source Location/Qualifiers  
1..5711  
BASE COUNT 1953 a 1099 c 1277 g 1382 t  
ORIGIN

Query Match 100.0%; Score 201; DB 6; Length 5711;  
Best Local Similarity 100.0%; Pred. No. 5.3e-103;  
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AGGCACGACGAGAACTCAACACTCATGAGTAAGAACTGACCTGGAGCCAAAGAG 60  
|||||  
Db 2100 AGGCACGACGAGAACTCAACACTCATGAGTAAGAACTGACCTGGAGCCAAAGAG 2159  
Oy 61 AGTAACAAGCCCAATGAGACAGACAGATAAAGACATGACAGTACTTCCAGAGCTG 120  
|||||  
Db 2160 AGTAACAAGCCCAATGAGACAGACAGATAAAGACATGACAGTACTTCCAGAGCTG 2219  
Oy 121 AAGTTACAATGCACTGGTCTTTTACTAAGTTCATAATACAGTGAATTAAAGAA 180  
|||||  
Db 2220 AAGTTACAATGCACTGGTCTTTTACTAAGTTCATAATACAGTGAATTAAAGAA 2279  
Oy 181 TTGTCAATCCTAGCCTTCCA 201  
|||||  
Db 2280 TTGTCAATCCTAGCCTTCCA 2300

RESULT 6  
AF110102 298 bp DNA linear PRI 08-JUN-1999  
LOCUS Homo sapiens breast and ovarian cancer susceptibility gene (BRCA1),  
DEFINITION exon 11 sequence.  
ACCESSION AF110102  
VERSION AF110102.1 GI:5006431  
KEYWORDS  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.  
REFERENCE 1 (bases 1 to 298)  
AUTHORS Corvello,C.M., Duarte,A.P.M., Mourao-Neto,M. and Simpson,A.J.G.  
TITLE Germ-line mutation at exon 11 of the BRCA1 gene in a female with  
family history of breast cancer  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 298)  
AUTHORS Corvello,C.M., Duarte,A.P.M., Monteiro,E. and Simpson,A.J.G.  
TITLE Direct Submission  
JOURNAL Submitted (30-NOV-1998) Cancer Genetics, Ludwig Institute for  
Cancer Research, R. Prof Antonio Prudente, 105 - 4 andar, Sao Paulo  
01509-010, Brazil  
Location/Qualifiers

source 1..298  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="17"  
/map="17q21"  
/cell\_type="Lymphocytes"  
<1..>298  
/gene="BRCA1"  
<1..>298  
/gene="BRCA1"  
/number=11  
misc\_feature 1..298  
/gene="BRCA1"  
/note="corresponds to breast and ovarian cancer  
susceptibility protein that is prematurely truncated due  
to mutation"  
55  
variation 55  
/gene="BRCA1"  
/replace="g"  
BASE COUNT 119 a 60 c 56 g 63 t  
ORIGIN

Query Match 74.6%; Score 150; DB 9; Length 298;  
Best Local Similarity 99.5%; Pred. No. 5.8e-74;  
Matches 200; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 AGGCACGACGAGAACTCAACACTCATGAGTAAGAACTGACCTGGAGCCAAAGAG 60  
|||||  
Db 91 AGGCACGACGAGAACTCAACACTCATGAGTAAGAACTGACCTGGAGCCAAAGAG 150  
Oy 61 AGTAACAAGCCCAATGAGACAGACAGATAAAGACATGACAGTACTTCCAGAGCTG 120  
|||||  
Db 151 AGTAACAAGCCCAATGAGACAGACAGATAAAGACATGACAGTACTTCCAGAGCTG 210  
Oy 121 AAGTTACAATGCACTGGTCTTTTACTAAGTTCATAATACAGTGAATTAAAGAA 180  
|||||  
Db 211 AAGTTACAATGCACTGGTCTTTTACTAAGTTCATAATACAGTGAATTAAAGAA 270  
Oy 181 TTGTCAATCCTAGCCTTCCA 201  
|||||  
Db 271 TTGTCAATCCTAGCCTTCCA 291

RESULT 7  
AF274503 649 bp mRNA linear PRI 06-JUN-2002  
LOCUS Homo sapiens breast and ovarian cancer susceptibility (BRCA1)  
DEFINITION pseudogene, partial mRNA sequence.  
ACCESSION AF274503  
VERSION AF274503.1 GI:8745572  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 649)  
AUTHORS Yassae,V.R., Zeinali,S., Harlirchi,I., Jarvandi,S.,  
Monaeghi,M.A., Hornby,D.P. and Dalton,A.  
TITLE Novel mutations in the BRCA1 and BRCA2 genes in Iranian women with  
early-onset breast cancer  
JOURNAL Breast Cancer Res. 4 (4), R6 (2002)  
MEDLINE 220944890  
PUBMED 12100744  
REFERENCE 2 (bases 1 to 649)  
AUTHORS Yassae,V.R.  
TITLE Direct Submission  
JOURNAL Submitted (02-JUN-2000) Molecular Biology and Biotechnology, The  
University of Sheffield, D27, Firth Court, Western Bank, Sheffield,  
South Yorkshire S10 2TN, UK  
Location/Qualifiers  
1..649  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

gene

<1..>649  
/gene="BRCA1"  
/note="breast and ovarian cancer susceptibility; mutant  
form; similar to Homo sapiens BRCA1 in Genbank Accession  
Number U14680"  
/pseudo

BASE COUNT 253 a 112 c 141 g 143 t  
ORIGIN

Query Match 74.6%; Score 150; DB 9; Length 649;  
Best Local Similarity 99.5%; Pred. No. 5.1e-74;  
Matches 200; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGGCACAGCAGAAACCTACACTCATGAGGTAAGAACCTGCAACTGGAGCCAGAAG 60  
DB 85 AGGCACAGCAGAAACCTACACTCATGAGGTAAGAACCTGCAACTGGAGCCAGAAG 144  
QY 61 AGTAACACCCCAATATACAGACAGCAAGTAAAGACATGACAGTATCTTTCCAGAGCTG 120  
DB 145 AGTAACACCCCAATATACAGACAGCAAGTAAAGACATGACAGTATCTTTCCAGAGCTG 204  
QY 121 AAGTTAACAAATGACACCTGTTCTTTACTAGTGTTCATATACAGTAACTTAAAGAA 180  
DB 205 AAGTTAACAAATGACACCTGTTCTTTACTAGTGTTCATATACAGTAACTTAAAGAA 264  
QY 181 TTTGTCAATCTAGCCTTCCA 201  
DB 265 TTTGTCAATCTAGCCTTCCA 285

RESULT 8  
AR004691 4249 bp DNA linear PAT 04-DEC-1998  
LOCUS Sequence 21 from patent US 5747282.  
DEFINITION AR004691  
ACCESSION AR004691.1 GI:3965570  
VERSION  
KEYWORDS  
ORGANISM Unknown.  
SOURCE Unclassified.

REFERENCE 1 (bases 1 to 4249)  
AUTHORS Skolnick,M.H., Goldgar,D.E., Miki,Y., Swenson,J., Kamb,A.,  
Harshman,K.D., Shattuck-Eidens,D.M., Tavtigian,S.V., Wiseman,R.W.  
and Futreal,P. Andrew.  
17Q-linked breast and ovarian cancer susceptibility gene  
Patent: US 5747282-A 21 05-MAY-1998;  
LOCATION/Qualifiers  
source 1..4249  
/organism="unknown"

BASE COUNT 1480 a 739 c 908 g 1122 t  
ORIGIN

Query Match 74.6%; Score 150; DB 6; Length 4249;  
Best Local Similarity 99.5%; Pred. No. 5.1e-74;  
Matches 200; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGGCACAGCAGAAACCTACACTCATGAGGTAAGAACCTGCAACTGGAGCCAGAAG 60  
DB 1512 AGGCACAGCAGAAACCTACACTCATGAGGTAAGAACCTGCAACTGGAGCCAGAAG 1571  
QY 61 AGTAACACCCCAATATACAGACAGCAAGTAAAGACATGACAGTATCTTTCCAGAGCTG 120  
DB 1572 AGTAACACCCCAATATACAGACAGCAAGTAAAGACATGACAGTATCTTTCCAGAGCTG 1631  
QY 121 AAGTTAACAAATGACACCTGTTCTTTACTAGTGTTCATATACAGTAACTTAAAGAA 180  
DB 1632 AAGTTAACAAATGACACCTGTTCTTTACTAGTGTTCATATACAGTAACTTAAAGAA 1691  
QY 181 TTTGTCAATCTAGCCTTCCA 201  
DB 1692 TTTGTCAATCTAGCCTTCCA 1712

RESULT 9  
AR008177 4249 bp DNA linear PAT 04-DEC-1998  
LOCUS Sequence 21 from patent US 5753441.  
DEFINITION AR008177  
ACCESSION AR008177  
VERSION AR008177.1 GI:3967286  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 4249)  
AUTHORS Skolnick,M.H., Goldgar,D.E., Miki,Y., Swenson,J., Kamb,A.,  
Harshman,K.D., Shattuck-Eidens,D.M., Tavtigian,S.V., Wiseman,R.W.  
and Futreal,P. Andrew.  
17Q-linked breast and ovarian cancer susceptibility gene  
Patent: US 5753441-A 21 19-MAY-1998;  
LOCATION/Qualifiers  
source 1..4249  
/organism="unknown"

BASE COUNT 1480 a 739 c 908 g 1122 t  
ORIGIN

Query Match 74.6%; Score 150; DB 6; Length 4249;  
Best Local Similarity 99.5%; Pred. No. 5.1e-74;  
Matches 200; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGGCACAGCAGAAACCTACACTCATGAGGTAAGAACCTGCAACTGGAGCCAGAAG 60  
DB 1512 AGGCACAGCAGAAACCTACACTCATGAGGTAAGAACCTGCAACTGGAGCCAGAAG 1571  
QY 61 AGTAACACCCCAATATACAGACAGCAAGTAAAGACATGACAGTATCTTTCCAGAGCTG 120  
DB 1572 AGTAACACCCCAATATACAGACAGCAAGTAAAGACATGACAGTATCTTTCCAGAGCTG 1631  
QY 121 AAGTTAACAAATGACACCTGTTCTTTACTAGTGTTCATATACAGTAACTTAAAGAA 180  
DB 1632 AAGTTAACAAATGACACCTGTTCTTTACTAGTGTTCATATACAGTAACTTAAAGAA 1691  
QY 181 TTTGTCAATCTAGCCTTCCA 201  
DB 1692 TTTGTCAATCTAGCCTTCCA 1712

RESULT 10  
AR136960 4249 bp DNA linear PAT 16-JUN-2001  
LOCUS Sequence 21 from patent US 6162897.  
DEFINITION AR136960  
ACCESSION AR136960  
VERSION AR136960.1 GI:14478210  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 4249)  
AUTHORS Skolnick,M.H., Goldgar,D.E., Miki,Y., Swenson,J., Kamb,A.,  
Harshman,K.D., Shattuck-Eidens,D.M., Tavtigian,S.V., Wiseman,R.W.  
and Futreal,P. Andrew.  
17Q-linked breast and ovarian cancer susceptibility gene  
Patent: US 6162897-A 21 19-DEC-2000;  
LOCATION/Qualifiers  
source 1..4249  
/organism="unknown"

BASE COUNT 1480 a 739 c 908 g 1122 t  
ORIGIN

Query Match 74.6%; Score 150; DB 6; Length 4249;  
Best Local Similarity 99.5%; Pred. No. 5.1e-74;  
Matches 200; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGGCACAGCAGAAACCTACACTCATGAGGTAAGAACCTGCAACTGGAGCCAGAAG 60  
DB 1512 AGGCACAGCAGAAACCTACACTCATGAGGTAAGAACCTGCAACTGGAGCCAGAAG 1571

Oy	61	AGTAAACAAGCCAAATGAAACAGACAGTAAAGACATGACATGATCTTCTCCAGAGCTG	120
Db	1572	AGTAAACAAGCCAAATGAAACAGACAGTAAAGACATGACATGATCTTCTCCAGAGCTG	1631
Oy	121	AAGTTAACAAATGACACCTGTTCTTTTACTAAGTTCAAATACCAAGTGAACCTAAAGAA	180
Db	1632	AAGTTAACAAATGACACCTGTTCTTTTACTAAGTTCAAATACCAAGTGAACCTAAAGAA	1691
Oy	181	TTTGTCAAATCCTAGCCTTCCA	201
Db	1692	TTTGTCAAATCCTAGCCTTCCA	1712
RESULT	11		
LOCUS	I76961	4249 bp	DNA
DEFINITION	Sequence 21 from patent US 5693473.	linear	PAT 03-APR-1998
ACCESSION	I76961		
VERSION	I76961.1	GI:3013115	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 4249) Shattuck-Eidens,D.M., Simard,J., Durocher,F., Emi,M. and Nakamura,Y.		
TITLE	Linked breast and ovarian cancer susceptibility gene		
JOURNAL	Patent: US 5693473-A 21 02-DEC-1997;		
FEATURES	Location/Qualifiers		
SOURCE	1..4249		
	/organism="unknown"		
BASE COUNT	1480 a	739 c	908 g
ORIGIN			1122 t

Query Match		74.6%: Score 150; DB 6; Length 4249;
Best Local Similarity 99.5%: Pred. No. 5.1e-74;		
Matches	200; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
QY	1 AGGCACAGCAGAAACCTACAACTCATGGAAGAGTAAAGAAACCTGCACACTGAGAGCCAGAG 60	
Db	1512 AGGCACAGCAGAAACCTACAACTCATGGAAGAGTAAAGAAACCTGCACACTGAGAGCCAGAGAG 15711	
QY	61 AGTAAACAAGCCAAATGAAACAGACAAGTAAAGACATGACAGTATCTTTCCAGAGCTG 120	
Db	1572 AGTAAACAAGCCAAATGAAACAGACAAGTAAAGACATGACAGCGATCTTTCCAGAGCTG 16313	
QY	121 AAGTTAACAATGAGACCGGTTCTTTACTAGTGTTCAAATACCAGAGAACTTAAAGAA 180	
Db	1632 AAGTTAACAATGAGACCGCTGTTCTTTACTAGTGTTCAAATACCAGAGAACTTAAAGAA 1691	
QY	181 TTTGTCAATCCTAGCCTTCCA 201	
Db	1692 TTTGTCAATCCTAGCCTTCCA 1712	
RESULT	12	
LOCUS	180956	4249 bp DNA linear PAT 10-JUN-1998
DEFINITION	Sequence 21 from patent US 5709999.	
ACCESSION	180956	
VERSION	180956.1 GI:3209246	
KEYWORDS	Unknown.	
SOURCE	Unknown.	
ORGANISM	Unclassified.	
REFERENCE	1 (bases 1 to 4249)	
AUTHORS	Shattuck-Eidens,D.M., Simard,J., Durocher,F., Eml,M. and Nakamura,Y.	
TITLE	Linked breast and ovarian cancer susceptibility gene	
JOURNAL	Patent: US 5709999-A 21 20-JUN-1998;	
FEATURES	location/qualifiers	
source	1..4249	
	/organism="unknown"	

ORIGIN	BASE COUNT	1480 a	739 c	908 g	1122 t
Query Match	74.6%;				
Best Local Similarity	99.5%;				
Matches	200; Conservative	0;	Mismatches	1; Indels	0; Gaps
Oy	1	AGGCACACCAAAACCTACACTCATGTGAAGAGTAAAGAACTGCACTGGAGCCAAAGAG	60		
Db	1512	AGGCACACCAAAACCTACACTCATGTGAAGAGTAAAGAACTGCAACTGGAGCCAAAGAG	1517		
Oy	61	AGTAACAAGCCAAATGAAACAGACAAGTAAAGACATGCAGTGTACTTTCCAGAGCTG	120		
Db	1572	AGTAACAAGCCAAATGAAACAGACAAGTAAAGACATGCAGTGTACTTTCCAGAGCTG	1631		
Oy	121	AAGTTAACAAATGCACCTGGTCTTTTACTAAGTGTTCAAATACCAGTGAACCTTAAAGAA	180		
Db	1632	AAGTTAACAAATGCACCTGGTCTTTTACTAAGTGTTCAAATACCAGTGAACCTTAAAGAA	1691		
Oy	181	TTTGTCAATCCTAGCCTTCCA	201		
Db	1692	TTTGTCAATCCTAGCCTTCCA	1712		

RESULT 13									
LOCUS	181052	181052	4249 bp	DNA	linear	PAT 10-JUN-1998			
DEFINITION	Sequence 21 from patent US 5710001.								
ACCESSION	181052								
VERSION	181052.1	GI:3209342							
KEYWORDS									
SOURCE	Unknown.								
ORGANISM	Unknown.								
REFERENCE	Unclassified.								
AUTHORS	1 (bases 1 to 4249) Skolnick,M.H., Goldgar,D.E., Miki,Y., Swenson,J., Kamb,A., Hershman,K.D., Shattuck-Eidens,D.M., Tavtigian,S.V., Wiseman,R.W. and Futreal,P.Andrew. 1/9-linked breast and ovarian cancer susceptibility gene Patent: US 5710001-A 21 20-JAN-1998; Location/Qualifiers 1..4249 /organism="unknown"								
TITLE									
JOURNAL									
FEATURES	source								
BASE COUNT	1480 a 739 c 908 g 1122 t								
ORIGIN									
Query Match	74.6%: Score 150; DB 6; Length 4249;								
Best Local Similarity	99.5%: Pred. No. 5.1e-74;								
Matches 200; Conservative	0; Mismatches 1; Indels 0; Gaps 0;								
QY	1 AGGCACAGCAGAAACTACACTCATCATGTGAAAGTAAAGAACTGCACACTGGAGCCAGNAG 60								
Db	1512 AGGCACAGCAGCAAAACTACACTCATCATGTGAAAGTAAAGAACTGCACACTGGAGCCAGNAG 1571								
QY	61 AGTAAACAAGCCAAATGAACAGACAGCAAGTAAAGACATGACAGTGTACTTTCCAGACCTG 120								
Db	1572 AGTAAACAAGCCAAATGAACAGACAGCAAGTAAAGACATGACAGCGTACTTTCCAGACCTG 1631								
QY	121 AAGTAAACAATGACACCTGGTCTTTTACTAAGTGTCAATATACAGTGAAGCTTAAAGAA 180								
Db	1632 AAGTAAACAATGACACCTGGTCTTTTACTAAGTGTCAATATACAGTGAAGCTTAAAGAA 1691								
QY	181 TTGTCAATCCTAGCCTTCCA 201								
Db	1692 TTGTCAATCCTAGCCTTCCA 1712								
RESULT 14									
LOCUS	AR048657	5656 bp	DNA	linear	PAT 29-SEP-1999				
DEFINITION	Sequence 1 from patent US 5821328.								
ACCESSION	AR048657								
VERSION	AR048657.1	GI:5971000							

KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
AUTHORS 1 (bases 1 to 5656)  
TITLE King, M.-C., Friedman, L., Ostermeyer, B., Rowell, S., Lynch, E., Szabo, C., and Lee, M.  
JOURNAL Genetic markers for breast, ovarian, and prostatic cancer  
FEATURES Patent: US 5821328-A 1 13-OCT-1998;  
Location/Qualifiers  
1.5656  
Source /organism="unknown"  
BASE COUNT 1939 a 1087 c 1263 g 1367 t  
ORIGIN

Query Match 74.6%: Score 150; DB 6; Length 5656;  
Best Local Similarity 99.5%: Pred. No. 5.1e-74;  
Matches 200; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGGCACAGCAGAACTTACACTGATGAGGTAAGAAGCTGCAACTGGAGCCAGAAG 60  
DB 2046 AGGCACAGCAGAACTTACACTGATGAGGTAAGAAGCTGCAACTGGAGCCAGAAG 2105  
QY 61 AGTAACAAGCCAAATGAGACAGACAGACATGACATGATGATCTTTCCAGAGCTG 120  
DB 2106 AGTAACAAGCCAAATGAGACAGACAGACATGACATGATGATCTTTCCAGAGCTG 2165  
QY 121 AAGTTAACAAATGACCTGGTCTTTTACTAGTGTCAATACAGTGAAGTAAAGAA 180  
DB 2166 AAGTTAACAAATGACCTGGTCTTTTACTAGTGTCAATACAGTGAAGTAAAGAA 2225  
QY 181 TTTGTCAATCCTAGCCTTCCA 201  
DB 2226 TTTGTCAATCCTAGCCTTCCA 2246

RESULT 15  
LOCUS 140792 5656 bp DNA linear PAT 13-MAY-1997  
DEFINITION Sequence 1 from patent US 5622829.  
ACCESSION 140792  
VERSION 140792.1 GI:2082272  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
AUTHORS 1 (bases 1 to 5656)  
TITLE King, M.-C., Friedman, L., Ostermeyer, B., Rowell, S., Lynch, E., Szabo, C., and Lee, M.  
JOURNAL Genetic markers for breast, ovarian, and prostatic cancer  
FEATURES Patent: US 5622829-A 1 22-APR-1997;  
Location/Qualifiers  
1.5656  
Source /organism="unknown"  
BASE COUNT 1939 a 1087 c 1263 g 1367 t  
ORIGIN

Query Match 74.6%: Score 150; DB 6; Length 5656;  
Best Local Similarity 99.5%: Pred. No. 5.1e-74;  
Matches 200; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGGCACAGCAGAACTTACACTGATGAGGTAAGAAGCTGCAACTGGAGCCAGAAG 60  
DB 2046 AGGCACAGCAGAACTTACACTGATGAGGTAAGAAGCTGCAACTGGAGCCAGAAG 2105  
QY 61 AGTAACAAGCCAAATGAGACAGACAGACATGACATGATGATCTTTCCAGAGCTG 120  
DB 2106 AGTAACAAGCCAAATGAGACAGACAGACATGACATGATGATCTTTCCAGAGCTG 2165  
QY 121 AAGTTAACAAATGACCTGGTCTTTTACTAGTGTCAATACAGTGAAGTAAAGAA 180  
DB 2166 AAGTTAACAAATGACCTGGTCTTTTACTAGTGTCAATACAGTGAAGTAAAGAA 2225

QY 181 TTTGTCAATCCTAGCCTTCCA 201  
DB 2226 TTTGTCAATCCTAGCCTTCCA 2246

Search completed: June 27, 2003, 06:01:23  
Job time: 1929 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

```
Run on:      June 26, 2003, 16:36:04 ; Search time 241 Seconds
              (without alignments)
              1878.223 Million cell updates/sec
```

Title: US-09-734-672-3\_COPY\_2100\_2300  
Perfect score: 301

Sequence: 1 AGGCACAGCAGAACTTACA.....TTGTCATCTTAGCTTCCA 201

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 1125999159 residues

```
word size : 0
```

Total number of hits satisfying chosen parameters: 4370478

Minimum	DB	seq	length:	0
Maximum	DB	seq	length:	3

Maximum DB seq length: 20000000000

Post-processing: Listing first 45 summaries

Database : N\_Geneseq\_101002:\*

1:	/SID5/gcgdata/genseq/genseqn-emb1/NA1980.DAT.*
2:	/SID5/gcgdata/genseq/genseqn-emb1/NA1981.DAT.*
3:	/SID5/gcgdata/genseq/genseqn-emb1/NA1982.DAT.*
4:	/SID52/gcgdata/genseq/genseqn-emb1/NA1983.DAT.*
5:	/SID52/gcgdata/genseq/genseqn-emb1/NA1984.DAT.*
6:	/SID52/gcgdata/genseq/genseqn-emb1/NA1985.DAT.*
7:	/SID52/gcgdata/genseq/genseqn-emb1/NA1986.DAT.*
8:	/SID52/gcgdata/genseq/genseqn-emb1/NA1987.DAT.*
9:	/SID52/gcgdata/genseq/genseqn-emb1/NA1988.DAT.*
10:	/SID52/gcgdata/genseq/genseqn-emb1/NA1989.DAT.*
11:	/SID52/gcgdata/genseq/genseqn-emb1/NA1990.DAT.*
12:	/SID52/gcgdata/genseq/genseqn-emb1/NA1991.DAT.*
13:	/SID52/gcgdata/genseq/genseqn-emb1/NA1992.DAT.*
14:	/SID52/gcgdata/genseq/genseqn-emb1/NA1993.DAT.*
15:	/SID52/gcgdata/genseq/genseqn-emb1/NA1994.DAT.*
16:	/SID52/gcgdata/genseq/genseqn-emb1/NA1995.DAT.*
17:	/SID52/gcgdata/genseq/genseqn-emb1/NA1996.DAT.*
18:	/SID52/gcgdata/genseq/genseqn-emb1/NA1997.DAT.*
19:	/SID52/gcgdata/genseq/genseqn-emb1/NA1998.DAT.*
20:	/SID52/gcgdata/genseq/genseqn-emb1/NA1999.DAT.*
21:	/SID52/gcgdata/genseq/genseqn-emb1/NA2000.DAT.*
22:	/SID52/gcgdata/genseq/genseqn-emb1/NA2001A.DAT.*
23:	/SID52/gcgdata/genseq/genseqn-emb1/NA2001B.DAT.*
24:	/SID52/gcgdata/genseq/genseqn-emb1/NA2002.DAT.*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	201	100.0	5711	18	AA87085	Human BRCA1 gene c
2	201	100.0	5711	19	AA662180	BRCA1 (om11) codin
3	201	100.0	5711	19	AAV46466	Human BRCA1 om13 p
4	201	100.0	5711	19	AAV46467	Human BRCA1 om13 p
5	201	100.0	5711	19	AAV46468	Human BRCA1 om13 p
6	201	100.0	5711	19	AAV46469	Human BRCA1 om13 p
7	201	100.0	5711	19	AAV46470	Human BRCA1 om13 p
8	201	100.0	5711	19	AAV46471	Human BRCA1 om13 p
9	201	100.0	5711	19	AAV46450	Human BRCA1 om13 c

10	201	100.0	5711	19	AAV64452	Human BRCA1 om11 F
11	201	100.0	5711	19	AAV64453	Human BRCA1 om11 F
12	201	100.0	5711	19	AAV64454	Human BRCA1 om11 F
13	201	100.0	5711	19	AAV64455	Human BRCA1 om11 F
14	201	100.0	5711	19	AAV64456	Human BRCA1 om11 F
15	201	100.0	5711	19	AAV64457	Human BRCA1 om11 F
16	201	100.0	5711	19	AAV64448	Human BRCA1 om11 C
17	201	100.0	5711	21	AAC60793	Human BRCA1 (om11)
18	201	100.0	5711	21	AAC60794	Human BRCA1 (om12)
19	201	100.0	5914	17	AACT17495	Mutated BRCA1 cod1
20	150	74.6	5656	17	AAAT42021	BRCA1 allele #580321
21	150	74.6	5656	18	AAAT70064	Mutant BRCA1 allele
22	150	74.6	5656	19	AAV60560	Tumourogenic BRCA1
23	150	74.6	5689	17	AAAT42023	BRCA1 allele #9815
24	150	74.6	5689	18	AAAT70066	Mutant BRCA1 allele
25	150	74.6	5689	19	AAV60562	Tumourogenic BRCA1
26	150	74.6	5707	17	AAAT42031	BRCA1 allele #582
27	150	74.6	5707	18	AAAT70074	Mutant BRCA1 allele
28	150	74.6	5707	19	AAV60570	Tumourogenic BRCA1
29	150	74.6	5709	17	AAAT42027	BRCA1 allele #6401
30	150	74.6	5709	17	AAAT42028	BRCA1 allele #4406
31	150	74.6	5709	17	AAAT42029	BRCA1 allele #1020
32	150	74.6	5709	17	AAAT42022	BRCA1 allele #9601
33	150	74.6	5709	18	AAAT70065	Mutant BRCA1 allele
34	150	74.6	5709	18	AAAT70072	Mutant BRCA1 allele
35	150	74.6	5709	18	AAAT70071	Mutant BRCA1 allele
36	150	74.6	5709	19	AAAT70070	Mutant BRCA1 allele
37	150	74.6	5709	19	AAV60561	Tumourogenic BRCA1
38	150	74.6	5709	19	AAV60566	Tumourogenic BRCA1
39	150	74.6	5709	19	AAV60567	Tumourogenic BRCA1
40	150	74.6	5709	19	AAV60568	Tumourogenic BRCA1
41	150	74.6	5709	24	AAV119138	Human BRCA1 DNA 38
42	150	74.6	5709	24	AAAS9013	Human DNA sequence
43	150	74.6	5710	17	AAAT42026	BRCA1 allele #388
44	150	74.6	5710	18	AAAT70065	Mutant BRCA1 allele
45	150	74.6	5710	19	AAV60565	Tumourogenic BRCA1

## ALIGNMENTS

RESULT 1  
AAT87085  
ID AAT87085 standard; cDNA; 5711 BP.

AC AAT87085;

DT 06-JAN-1998 (first entry)

Human BRCA1 gene consensus

KM BRCA1 gene; BRCA1(omi); breast cancer; ovary cancer; polymorphism  
KM genetic testing; diagnosis; gene therapy; ss.

OS Homo sapiens

FH	Key	Location/Qualifiers
FE	CCC	120 7711

FT	/+tag=	b
FT	/note=	"AGC (40%) and AGT (55-65%) polymorphism
FT		at position 2201"
FT	variation	2430

FT	/note=	"35-45% TTG (Leu) and 55-65% CTG (Leu)
FT		polymorphism at position 2430"
FT	variation	
	2731	

FT	/note=	"35-35% CCG (Pro) and 65-75% CTG (Leu)
FT		polymorphism at position 2731"
FT	variation	
	3232	
FT	/*tag=	e

Query Match	Best Local Similarity	Matches	Conservative	Mismatches	Indels	Gaps
1	AGGCACAGCAAACTTACAACTGAGAGGTAAGAAAGCTGCAACTGAGCCAGAAG	60				
2100	AGGCACAGCAAACTTACAACTGAGAGGTAAGAAAGCTGCAACTGAGCCAGAAG	2159				
61	AGTAAACAAGCAAAATGAACAGACAAGTAAAGACATGACAGTAACTTCCAGAGCTG	120				
2160	AGTAAACAAGCAAAATGAACAGACAAGTAAAGACATGACAGTAACTTCCAGAGCTG	2219				
121	AACTTAACAAGTACAGACCTGCTTCTTTTACTAAGTGTTCAAATATACAGTGAAGCTTAAAGAA	180				
2220	AACTTAACAAGTACAGACCTGCTTCTTTTACTAAGTGTTCAAATATACAGTGAAGCTTAAAGAA	2279				
181	TTTGTCAATCTAGACCTTCCA	201				
2280	TTTGTCAATCTAGACCTTCCA	2300				

RESULT 2
AAV62180
ID AAV62180 standard; DNA: 5711 BP.
XX
AC AAV62180;
DT 11-FEB-1999 (first entry)
DE BRCA1 (oml1) coding sequence.
XX
KW BRCA1; mutation detection; disease screening; multiple allele variation; breast cancer; ovarian cancer; cystic fibrosis; Li-Fraumeni syndrome; Duchenne muscular dystrophy; Becker muscular dystrophy; ss.
KM
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 120..5711
FT /tag= a
PN WO9844157-A2.
XX
PD 08-OCT-1998.
PF 26-MAR-1998; 98MO-USO6002.
PR 28-MAR-1997; 97US-0825487.
XX
PA (ONCO-) ONCOMED INC.
PI Murphy PD, White MB;
XX
DR WPI: 1998-542713/46.
XX P-PSDB; AAM79665.
PT Identifying variations in polynucleotide sequences - using allele specific hybridisation assay, sequence variation locating assay, and direct sequencing, in a stepwise procedure
PS Disclosure: Fig la-j; 62pp; English.
XX
CC This sequence encodes the human BRCA (oml1) protein, and was used to test the method of the invention. The method is for determining the presence or absence of a sequence variation in a gene sample, and comprises: CC (a) performing an allele specific hybridisation assay for one or more pre-determined sequence variations; (b) if no pre-determined sequence variation found in step (a) then performing a sequence variation location assay; (c) if no sequence variation found in step (b) then sequencing the gene sample; (c1) if sequence variation is found in step (b) then targeted confirmatory sequencing is performed; and (d) determining the presence of a sequence variation by analysing the sequence(s) obtained in step (c1) or step (c1) against a reference sample. Alternatively, step (a) or step (b) is omitted from the method. The invention provides a stepwise and integrated method for the efficient and accurate detection of variations in polynucleotide sequences, being directed towards screening for diseases associated with multiple allele variations, including breast and ovarian cancer, cystic fibrosis, Duchenne and Becker muscular dystrophy, and Li-Fraumeni syndrome.
SO Sequence 5711 BP; 1953 A; 1099 C; 1277 G; 1382 T; 0 other:
Oy Query Match 100.0%; Score 201; DB 19; Length 5711; Best Local Similarity 100.0%; Pred. No. 6,4e-97; Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Dy 1 AGGACAGCAGGAAGACTTACATCTCATGTGAGAGTAAGACCTGACACTGAGCGAAGANG 60     Db 2100 AGGCACAGCAGGAAGACTTACATCTCATGTGAGAGTAAGAACCTGCACTGAGCGAAGANG 2155      Oy 61 AGTAGACAGCCAAATGACAGACAGTAGTAAAGACATGACAGTGATATCTTTCCAGAGCTG 120 

Db 2160 AGTAAACAGCCAAATGAACAGACAGTAAAGACATGACATGATCTTCCAGAGCTG 2219  
 QY 121 AGTTACCAATGACCTGGTCTTTACTAGTGTCAATACCACTGAACCTAAAGAA 180  
 Db 2220 AAGTTACCAATGACCTGGTCTTTACTAGTGTCAATACCACTGAACCTAAAGAA 2279  
 QY 181 TTGTCAATCCTAGCCTTCCA 201  
 Db 2280 TTGTCAATCCTAGCCTTCCA 2300

RESULT 3  
 AAV46466 ID AAV46466 standard; cDNA; 5711 BP.  
 AC AAV46466;  
 XX  
 DT 18-NOV-1998 (first entry)  
 XX  
 DE Human BRCA1 omi3 polymorphism #2 cDNA.  
 XX  
 KW BRCA1: omi3; human; breast and ovarian cancer predisposing gene;  
 KM polymorphism; susceptibility; anti-oncogene; tumour suppressor;  
 KM chromosome 17q; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 120..5711  
 FT /\*tag= a  
 FT /product= "BRCA1 omi3 protein"  
 FT 2430  
 FT variation  
 FT /\*tag= b  
 FT /note= "This polymorphic variation can be a T or C  
 nucleotide"  
 FT  
 FT  
 XX  
 XX US5750400-A.  
 PN 12-MAY-1998.  
 XX  
 PD 12-FEB-1997; 97US-0798691.  
 XX  
 PF 12-FEB-1996; 96US-0598591.  
 PR 12-FEB-1997; 97US-0798691.  
 PR XX  
 XX (ONCO-) ONCORMED INC.  
 PA Allen AC, Alvaras CP, Critz BS, Murphy PD, Olson SJ;  
 PI Schelter DB, Zeng B;  
 PI  
 XX  
 XX WPI; 1998-296774/26.  
 DR  
 XX  
 PT BRCA1 omi gene coding sequences - useful for distinguishing between  
 PT polymorphisms and mutation(s) in the screening for disposition to  
 PT breast or ovarian cancer  
 PT  
 XX  
 PS Claim 2e; Page -: 54pp; English.  
 XX  
 CC This sequence encodes a human BRCA1 (breast and ovarian cancer  
 CC predisposing gene) omi3 gene in which a polymorphic variation occurs at  
 CC nucleotide 2430. This sequence and other polymorphic variations of this  
 CC sequence are useful for the identification of an individual who may or  
 CC may not have an increased susceptibility to breast or ovarian cancer.  
 CC The sequences used identify gene changes which are due to polymorphisms,  
 CC rather than cancer causing mutations. BRCA1 is an anti-oncogene (tumour  
 CC suppressor) which is involved in genetic inheritance of cancers,  
 CC especially breast and ovarian cancer. It is found at human chromosome  
 CC 17q which is known to be linked to cancer susceptibility, especially  
 CC breast cancer. Cells containing a mutation in this gene lose the  
 CC wild-type function of BRCA1 and are more susceptible to cancers.  
 CC NOTE: This sequence does not appear in the specification but has been  
 CC created from the wild type BRCA1 omi3 gene represented in AAV46450.  
 CC  
 XX

SO Sequence 5711 BP; 1953 A; 1097 C; 1277 G; 1383 T; 1 other;  
 Query Match 100.0%; Score 201; DB 19; Length 5711;  
 Best Local Similarity 100.0%; Pred. No. 6,4e-97;  
 Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGCACAGCAGAAACCTACACATCATGAGGTTAAAGAACCTGACATGAGCCAAAG 60  
 Db 2100 AGGCACAGCAGAAACCTACACATCATGAGGTTAAAGAACCTGACATGAGCCAAAG 2159  
 QY 61 AGTAAACAGCCAAATGAACAGACAGTAAAGACATGACATGATCTTCCAGAGCTG 120  
 Db 2160 AGTAAACAGCCAAATGAACAGACAGTAAAGACATGACATGATCTTCCAGAGCTG 2219  
 QY 121 AAGTTACCAATGACCTGGTCTTTACTAGTGTCAATACCACTGAACCTAAAGAA 180  
 Db 2220 AAGTTACCAATGACCTGGTCTTTACTAGTGTCAATACCACTGAACCTAAAGAA 2279  
 QY 181 TTGTCAATCCTAGCCTTCCA 201  
 Db 2280 TTGTCAATCCTAGCCTTCCA 2300

RESULT 4  
 AAV46467 ID AAV46467 standard; cDNA; 5711 BP.  
 AC AAV46467;  
 XX  
 DT 18-NOV-1998 (first entry)  
 XX  
 DE Human BRCA1 omi3 polymorphism #3 cDNA.  
 XX  
 KW BRCA1: omi3; human; breast and ovarian cancer predisposing gene;  
 KM polymorphism; susceptibility; anti-oncogene; tumour suppressor;  
 KM chromosome 17q; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 120..5711  
 FT /\*tag= a  
 FT /product= "BRCA1 omi3 protein"  
 FT 2731  
 FT variation  
 FT /\*tag= b  
 FT /note= "This polymorphic variation can be a C or T  
 nucleotide"  
 FT  
 FT  
 XX  
 XX US5750400-A.  
 PN 12-MAY-1998.  
 XX  
 PD 12-FEB-1997; 97US-0798691.  
 XX  
 PF 12-FEB-1996; 96US-0598591.  
 PR 12-FEB-1997; 97US-0798691.  
 PR XX  
 XX (ONCO-) ONCORMED INC.  
 PA Allen AC, Alvaras CP, Critz BS, Murphy PD, Olson SJ;  
 PI Schelter DB, Zeng B;  
 PI  
 XX  
 XX WPI; 1998-296774/26.  
 DR  
 XX  
 PT BRCA1 omi gene coding sequences - useful for distinguishing between  
 PT polymorphisms and mutation(s) in the screening for disposition to  
 PT breast or ovarian cancer  
 PT  
 XX  
 PS Claim 2e; Page -: 54pp; English.  
 XX  
 CC This sequence encodes a human BRCA1 (breast and ovarian cancer  
 CC predisposing gene) omi3 gene in which a polymorphic variation occurs at  
 CC nucleotide 2731. This sequence and other polymorphic variations of this  
 CC



CC sequence are useful for the identification of an individual who may or  
CC may not have an increased susceptibility to breast or ovarian cancer.  
CC The sequences used identify gene changes which are due to polymorphisms,  
CC rather than cancer causing mutations. BRCA1 is an anti-oncogene (tumour  
CC suppressor) which is involved in genetic inheritance of cancers,  
CC especially breast and ovarian cancer. It is found at human chromosome  
CC 17q which is known to be linked to cancer susceptibility, especially  
CC breast cancer. Cells containing a mutation in this gene lose the  
CC wild-type function of BRCA1 and are more susceptible to cancers.  
CC NOTE: This sequence does not appear in the specification but has been  
CC created from the wild type BRCA1 omi3 gene represented in AAV46450.

CC Sequence 5711 BP; 1953 A; 1098 C; 1277 G; 1382 T; 1 other;

Query Match 100.0%; Score 201; DB 19; Length 5711;  
Best Local Similarity 100.0%; Pred. No. 6.4e-97;  
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AGGCACAGCAGAAACCTACACTGAGTAAGAACTGCAACTGAGCCAGAG 60  
DB 2100 AGGCACAGCAGAAACCTACACTGAGTAAGAACTGCAACTGAGCCAGAG 2159  
OY 61 AGTACAGCCAAATGAGACAGACAGTAAGACATGATGATCTTCCAGAGCTG 120  
DB 2160 AGTACAGCCAAATGAGACAGACAGTAAGACATGATGATCTTCCAGAGCTG 2219  
OY 121 AAGTTAACAAATGACACCTGCTCTTTACTAGTGTCAATACAGTGAAGTAAGAA 180  
DB 2220 AAGTTAACAAATGACACCTGCTCTTTACTAGTGTCAATACAGTGAAGTAAGAA 2279  
OY 181 TTTGTCAATCCTAGCCTTCCA 201  
DB 2280 TTTGTCAATCCTAGCCTTCCA 2300

## RESULT 5

AAV46468 standard; cDNA; 5711 BP.

AAV46468;

18-NOV-1998 (first entry)

Human BRCA1 omi3 polymorphism #4 cDNA.

BRCA1: omi3; human: breast and ovarian cancer predisposing gene;  
polymorphism: susceptibility; anti-oncogene; tumour suppressor;  
chromosome 17q; ss.

Homo sapiens.

Location/Qualifiers  
120..5711

/tag= a "BRCA1 omi3 protein"

/product=

3232

/tag= b  
/note= "This polymorphic variation can be an A or G  
nucleotide"

US5750400-A.

12-MAY-1998.

12-FEB-1997; 97US-0798691.

12-FEB-1996; 96US-0598591.

12-FEB-1997; 97US-0798691.

(ONCO-) ONCORMED INC.

Allen AC, Alvares CP, Ciltz BS, Murphy PD, Olson SJ;

Scheller DB, Zeng B;

XX WPI: 1998-296774/26.

XX BRCA1 omi3 gene coding sequences - useful for distinguishing between

PT polymorphisms and mutation(s) in the screening for disposition to

PS breast or ovarian cancer

PS Claim 2e; Page -: 54pp; English.

CC This sequence encodes a human BRCA1 (breast and ovarian cancer  
CC predisposing gene) omi3 gene in which a polymorphic variation occurs at  
CC nucleotide 3232. This sequence and other polymorphic variations of this  
CC sequence are useful for the identification of an individual who may or  
CC may not have an increased susceptibility to breast or ovarian cancer.  
CC The sequences used identify gene changes which are due to polymorphisms,  
CC rather than cancer causing mutations. BRCA1 is an anti-oncogene (tumour  
CC suppressor) which is involved in genetic inheritance of cancers,  
CC especially breast and ovarian cancer. It is found at human chromosome  
CC 17q which is known to be linked to cancer susceptibility, especially  
CC breast cancer. Cells containing a mutation in this gene lose the  
CC wild-type function of BRCA1 and are more susceptible to cancers.  
CC NOTE: This sequence does not appear in the specification but has been  
CC created from the wild type BRCA1 omi3 gene represented in AAV46450.

CC Sequence 5711 BP; 1953 A; 1098 C; 1276 G; 1383 T; 1 other;

Query Match 100.0%; Score 201; DB 19; Length 5711;  
Best Local Similarity 100.0%; Pred. No. 6.4e-97;  
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AGGCACAGCAGAAACCTACACTGAGTAAGAACTGCAACTGAGCCAGAG 60  
DB 2100 AGGCACAGCAGAAACCTACACTGAGTAAGAACTGCAACTGAGCCAGAG 2159  
OY 61 AGTACAGCCAAATGAGACAGACAGTAAGACATGATGATCTTCCAGAGCTG 120  
DB 2160 AGTACAGCCAAATGAGACAGACAGTAAGACATGATGATCTTCCAGAGCTG 2219  
OY 121 AAGTTAACAAATGACACCTGCTCTTTACTAGTGTCAATACAGTGAAGTAAGAA 180  
DB 2220 AAGTTAACAAATGACACCTGCTCTTTACTAGTGTCAATACAGTGAAGTAAGAA 2279  
OY 181 TTTGTCAATCCTAGCCTTCCA 201  
DB 2280 TTTGTCAATCCTAGCCTTCCA 2300

## RESULT 6

AAV46469 standard; cDNA; 5711 BP.

AAV46469;

18-NOV-1998 (first entry)

Human BRCA1 omi3 polymorphism #5 cDNA.

BRCA1: omi2; human: breast and ovarian cancer predisposing gene;  
polymorphism: susceptibility; anti-oncogene; tumour suppressor;  
chromosome 17q; ss.

Homo sapiens.

Location/Qualifiers  
120..5711

/tag= a "BRCA1 omi3 protein"

/product=

3667

/tag= b  
/note= "This polymorphic variation can be an A or G  
nucleotide"

US5750400-A.

```
XX 12-MAY-1998.
XX
XX 12-FEB-1997; 97US-0798691.
XX
XX 12-FEB-1996; 96US-0598591.
XX
XX 12-FEB-1997; 97US-0798691.
XX
XX (ONCO-) ONCORMED INC.
XX
XX Allen AC, Alvares CP, Critz BS, Murphy PD, Olson SJ;
XX Scheller DB, Zeng B;
XX WPI; 1998-296774/26.
XX
XX BRCA1 om1 gene coding sequences - useful for distinguishing between
XX polymorphisms and mutation(s) in the screening for disposition to
XX breast or ovarian cancer
XX
XX Claim 2e; Page -: 54pp; English.
XX
XX This sequence encodes a human BRCA1 (breast and ovarian cancer
XX predisposing gene) om13 gene in which a polymorphic variation occurs at
XX nucleotide 3667. This sequence and other polymorphic variations of this
XX sequence are useful for the identification of an individual who may or
XX may not have an increased susceptibility to breast or ovarian cancer.
XX The sequences used identify gene changes which are due to polymorphisms,
XX rather than cancer causing mutations. BRCA1 is an anti-oncogene (tumour
XX suppressor) which is involved in genetic inheritance of cancers,
XX especially breast and ovarian cancer. It is found at human chromosome
XX 17q which is known to be linked to cancer susceptibility, especially
XX breast cancer. Cells containing a mutation in this gene lose the
XX wild-type function of BRCA1 and are more susceptible to cancers.
XX NOTE: This sequence does not appear in the specification but has been
XX created from the wild type BRCA1 om13 gene represented in AAV46450.
XX
XX Sequence 5711 BP; 1953 A; 1098 C; 1276 G; 1383 T; 1 other;
XX
XX Query Match 100.0%; Score 201; DB 19; Length 5711;
XX Best Local Similarity 100.0%; Pred. No. 6.4e-97;
XX Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 AGGCACAGCAGAAACCTACACATCATGGAAGGTAAGAACCTGAGAGCCAAAG 60
XX 2100 AGGCACAGCAGAAACCTACACATCATGGAAGGTAAGAACCTGAGAGCCAAAG 2159
XX
XX 61 AGTAACAGCCCAATGACAGACAGTAAGAACATGATGATCTTCCAGAGCTG 120
XX 2160 AGTAACAGCCCAATGACAGACAGTAAGAACATGATGATCTTCCAGAGCTG 2219
XX
XX 121 AAGTTAAACAATGACACCTGTTCTTTACTAAGTGTCAATTCAGTAACTTAAGAA 180
XX 2220 AAGTTAAACAATGACACCTGTTCTTTACTAAGTGTCAATTCAGTAACTTAAGAA 2279
XX
XX 181 TTGTCAATCTAGCCTTCCA 201
XX 2280 TTGTCAATCTAGCCTTCCA 2300
XX
XX
XX RESULT 7
XX AAV46470
XX ID AAV46470 standard; cDNA; 5711 BP.
XX
XX AAV46470;
XX
XX 18-NOV-1998 (first entry)
XX
XX Human BRCA1 om13 polymorphism #6 cDNA.
XX
XX BRCA1: om13: human: breast and ovarian cancer predisposing gene;
XX polymorphism: susceptibility; anti-oncogene; tumour suppressor;
XX chromosome 17q; ss.
XX
```

```
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 120..5711
XX
XX /*tag= a
XX /product= "BRCA1 om13 protein"
XX variation 4427
XX /*tag= b
XX /note= "this polymorphic variation can be a T or C
XX nucleotide"
XX
XX US5750400-A.
XX
XX 12-MAY-1998.
XX
XX 12-FEB-1997; 97US-0798691.
XX
XX 12-FEB-1996; 96US-0598591.
XX
XX 12-FEB-1997; 97US-0798691.
XX
XX (ONCO-) ONCORMED INC.
XX
XX Allen AC, Alvares CP, Critz BS, Murphy PD, Olson SJ;
XX Scheller DB, Zeng B;
XX WPI; 1998-296774/26.
XX
XX BRCA1 om1 gene coding sequences - useful for distinguishing between
XX polymorphisms and mutation(s) in the screening for disposition to
XX breast or ovarian cancer
XX
XX Claim 2e; Page -: 54pp; English.
XX
XX This sequence encodes a human BRCA1 (breast and ovarian cancer
XX predisposing gene) om13 gene in which a polymorphic variation occurs at
XX nucleotide 4427. This sequence and other polymorphic variations of this
XX sequence are useful for the identification of an individual who may or
XX may not have an increased susceptibility to breast or ovarian cancer.
XX The sequences used identify gene changes which are due to polymorphisms,
XX rather than cancer causing mutations. BRCA1 is an anti-oncogene (tumour
XX suppressor) which is involved in genetic inheritance of cancers,
XX especially breast and ovarian cancer. It is found at human chromosome
XX 17q which is known to be linked to cancer susceptibility, especially
XX breast cancer. Cells containing a mutation in this gene lose the
XX wild-type function of BRCA1 and are more susceptible to cancers.
XX NOTE: This sequence does not appear in the specification but has been
XX created from the wild type BRCA1 om13 gene represented in AAV46450.
XX
XX Sequence 5711 BP; 1953 A; 1098 C; 1277 G; 1382 T; 1 other;
XX
XX Query Match 100.0%; Score 201; DB 19; Length 5711;
XX Best Local Similarity 100.0%; Pred. No. 6.4e-97;
XX Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 AGGCACAGCAGAAACCTACACATCATGGAAGGTAAGAACCTGAGAGCCAAAG 60
XX 2100 AGGCACAGCAGAAACCTACACATCATGGAAGGTAAGAACCTGAGAGCCAAAG 2159
XX
XX 61 AGTAACAGCCCAATGACAGACAGTAAGAACATGATGATCTTCCAGAGCTG 120
XX 2160 AGTAACAGCCCAATGACAGACAGTAAGAACATGATGATCTTCCAGAGCTG 2219
XX
XX 121 AAGTTAAACAATGACACCTGTTCTTTACTAAGTGTCAATTCAGTAACTTAAGAA 180
XX 2220 AAGTTAAACAATGACACCTGTTCTTTACTAAGTGTCAATTCAGTAACTTAAGAA 2279
XX
XX 181 TTGTCAATCTAGCCTTCCA 201
XX 2280 TTGTCAATCTAGCCTTCCA 2300
XX
XX
XX RESULT 8
XX AAV46471
```

```

ID AAV46471 standard; cDNA; 5711 BP.
XX
XX AAV46471;
AC
XX
XX 18-NOV-1998 (first entry)
DT
XX
XX Human BRCA1 omi3 polymorphism #7 cDNA.
DE
XX
XX BRCA1; omi3; human; breast and ovarian cancer predisposing gene;
KM polymorphism; susceptibility; anti-oncogene; tumour suppressor;
KW chromosome 17q; ss.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH 120..5711
FT CDS /tag= a
FT /product= "BRCA1 omi3 protein"
FT variation 4956
FT /tag= b
FT /note= "This polymorphic variation can be an A or G
FT nucleotide"
XX
XX US5750400-A.
XX
XX 12-MAY-1998.
XX
XX 12-FEB-1997; 97US-0798691.
XX
XX 12-FEB-1996; 96US-0598591.
XX 12-FEB-1997; 97US-0798691.
XX
XX (ONCO-) ONCORMED INC.
XX
XX Allen AC, Alvares CP, Critz BS, Murphy PD, Olson SJ;
PI Schelter DB, Zeng B;
XX
XX WPI: 1998-296774/26.
XX
XX BRCA1 omi gene coding sequences - useful for distinguishing between
PT polymorphisms and mutation(s) in the screening for disposition to
PT breast or ovarian cancer
XX
XX Claim 2e; Page -: 54pp; English.
XX
XX This sequence encodes a human BRCA1 (breast and ovarian cancer
XX predisposing gene) omi3 gene in which a polymorphic variation occurs at
XX nucleotide 4956. This sequence and other polymorphic variations of this
XX sequence are useful for the identification of an individual who may or
XX may not have an increased susceptibility to breast or ovarian cancer.
XX The sequences used identify gene changes which are due to polymorphisms,
XX rather than cancer causing mutations. BRCA1 is an anti-oncogene (tumour
XX suppressor) which is involved in genetic inheritance of cancers,
XX especially breast and ovarian cancer. It is found at human chromosome
XX 17q which is known to be linked to cancer susceptibility, especially
XX breast cancer. Cells containing a mutation in this gene lose the
XX wild-type function of BRCA1 and are more susceptible to cancers.
XX NOTE: This sequence does not appear in the specification but has been
XX created from the wild type BRCA1 omi3 gene represented in AAV46450.
XX
XX Sequence 5711 BP; 1953 A; 1098 C; 1276 G; 1383 T; 1 other:
SQ
Query Match 100.0%; Score 201; DB 19; Length 5711;
Best Local Similarity 100.0%; Pred. No. 6,4e-97;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 AGGCACAGCAGAACTACACTCATGAGGTAAGAACTGCACTGGAGCCAGAG 60
Db 2100 AGGCACAGCAGAACTACACTCATGAGGTAAGAACTGCACTGGAGCCAGAG 2159
OY 61 AGTACAGCCCAATGAGACAGACAGTAAAGACATGACATGATCTTCCAGAGCTG 120
Db 2160 AGTACAGCCCAATGAGACAGACAGTAAAGACATGACATGATCTTCCAGAGCTG 2219

```

```

OY 121 AGCTAACAAATGCACCTGGTTCTTTTACTAAGTGTCAAAATACCACTGAATAAGAA 180
Db 2220 AGCTAACAAATGCACCTGGTTCTTTTACTAAGTGTCAAAATACCACTGAATAAGAA 2279
OY 181 TTGTGCAATCCTTAGCCTTCCA 201
Db 2280 TTGTGCAATCCTTAGCCTTCCA 2300
RESULT 9
ID AAV46450 standard; cDNA; 5711 BP.
XX
XX AAV46450;
AC
XX
XX 18-NOV-1998 (first entry)
DT
XX
XX Human BRCA1 omi3 cDNA.
DE
XX
XX BRCA1; omi3; human; breast and ovarian cancer predisposing gene;
KM polymorphism; susceptibility; anti-oncogene; tumour suppressor;
KW chromosome 17q; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH 120..5711
FT CDS /tag= a
FT /product= "BRCA1 omi3 protein"
XX
XX US5750400-A.
XX
XX 12-MAY-1998.
XX
XX 12-FEB-1997; 97US-0798691.
XX
XX 12-FEB-1996; 96US-0598591.
XX 12-FEB-1997; 97US-0798691.
XX
XX (ONCO-) ONCORMED INC.
XX
XX Allen AC, Alvares CP, Critz BS, Murphy PD, Olson SJ;
PI Schelter DB, Zeng B;
XX
XX WPI: 1998-296774/26.
XX
XX BRCA1 omi gene coding sequences - useful for distinguishing between
PT polymorphisms and mutation(s) in the screening for disposition to
PT breast or ovarian cancer
XX
XX Claim 2e; Column 55-62; 54pp; English.
XX
XX This sequence encodes the human BRCA1 (breast and ovarian cancer
XX predisposing gene) omi3 gene. This sequence and polymorphic variations of
XX this sequence are useful for the identification of an individual who may
XX or may not have an increased susceptibility to breast or ovarian cancer.
XX The sequences used identify gene changes which are due to polymorphisms,
XX rather than cancer causing mutations. BRCA1 is an anti-oncogene (tumour
XX suppressor) which is involved in genetic inheritance of cancers,
XX especially breast and ovarian cancer. It is found at human chromosome 17q
XX which is known to be linked to cancer susceptibility, especially breast
XX cancer. Cells containing a mutation in this gene lose the wild-type
XX function of BRCA1 and are more susceptible to cancers.
XX
XX Sequence 5711 BP; 1953 A; 1098 C; 1277 G; 1383 T; 0 other:
SQ
Query Match 100.0%; Score 201; DB 19; Length 5711;
Best Local Similarity 100.0%; Pred. No. 6,4e-97;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 AGGCACAGCAGAACTACACTCATGAGGTAAGAACTGCACTGGAGCCAGAG 60

```

```
Db 2100 AGGCACAGCAGAAACCTACACACTCATGGAAGTAAAGAACTGCACTGGAGCCAGAGAG 2159
Qy 61 AGTACACAGCCAAATGACAGACAGATAAAGACATGACACTGACTTTCCAGAGCTG 120
Db 2160 AGTACACAGCCAAATGACAGACAGATAAAGACATGACACTGACTTTCCAGAGCTG 2219
Qy 121 AAGTTACAAATGACACCTGGTTCTTTACTAGTGTCAATATACAGTGAACCTTAAGAA 180
Db 2220 AAGTTACAAATGACACCTGGTTCTTTACTAGTGTCAATATACAGTGAACCTTAAGAA 2279
Qy 181 TTGTCAATCCTTACGCTTCCA 201
Db 2280 TTGTCAATCCTTACGCTTCCA 2300

RESULT 10
AAV46452 ID AAV46452 standard; cDNA; 5711 BP.
AC AAV46452;
XX
DT 18-NOV-1998 (first entry)
XX
DE Human BRCA1 om1 polymorphism #2 cDNA.
XX
KW BRCA1: om1; human: breast and ovarian cancer predisposing gene;
KW polymorphism; susceptibility; anti-oncogene; tumour suppressor;
KW chromosome 17q; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 120..5711
FT FT /*tag= a
FT FT /*product= "BRCA1 om1 protein"
FT FT 2430
FT FT variation
FT FT /*tag= b
FT FT /*note= "This polymorphic variation can be a T or C
FT FT nucleotide"
XX
PN US5750400-A.
XX
PD 12-MAY-1998.
XX
PE 12-FEB-1997; 97US-0798691.
XX
PR 12-FEB-1996; 96US-0598591.
XX PR 12-FEB-1997; 97US-0798691.
XX
PA (ONCO-) ONCORMED INC.
XX
PI Allen AC, Alvares CP, Critz BS, Murphy PD, Olson SJ;
PI Schelter DB, Zeng B;
XX
DR WPI: 1998-296774/26.
XX
XX
PT BRCA1 om1 gene coding sequences - useful for distinguishing between
PT polymorphisms and mutation(s) in the screening for disposition to
PT breast or ovarian cancer
XX
PS Claim 2e: Page -: 54pp; English.
XX
XX
CC This sequence encodes a human BRCA1 (breast and ovarian cancer
CC predisposing gene) om1 gene in which a polymorphic variation occurs at
CC nucleotide 2430. This sequence and other polymorphic variations of this
CC sequence are useful for the identification of an individual who may or
CC may not have an increased susceptibility to breast or ovarian cancer.
CC The sequences used identify gene changes which are due to polymorphisms,
CC rather than cancer causing mutations. BRCA1 is an anti-oncogene (tumour
CC suppressor) which is involved in genetic inheritance of cancers,
CC especially breast and ovarian cancer. It is found at human chromosome
CC 17q which is known to be linked to cancer susceptibility, especially
```

```
CC breast cancer. Cells containing a mutation in this gene lose the
CC wild-type function of BRCA1 and are more susceptible to cancers.
CC NOTE: This sequence does not appear in the specification but has been
CC created from the wild type BRCA1 om1 gene represented in AAV46448.
XX
SO Sequence 5711 BP; 1953 A; 1098 C; 1277 G; 1382 T; 1 other;

Query Match 100.0%; Score 201; DB 19; Length 5711;
Best Local Similarity 100.0%; Pred. NO. 6.4e-97;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGCACAGCAGAAACCTACACACTCATGGAAGTAAAGAACTGCACTGGAGCCAGAGAG 60
Db 2100 AGGCACAGCAGAAACCTACACACTCATGGAAGTAAAGAACTGCACTGGAGCCAGAGAG 2159
Qy 61 AGTACACAGCCAAATGACAGACAGATAAAGACATGACACTGACTTTCCAGAGCTG 120
Db 2160 AGTACACAGCCAAATGACAGACAGATAAAGACATGACACTGACTTTCCAGAGCTG 2219
Qy 121 AAGTTACAAATGACACCTGGTTCTTTACTAGTGTCAATATACAGTGAACCTTAAGAA 180
Db 2220 AAGTTACAAATGACACCTGGTTCTTTACTAGTGTCAATATACAGTGAACCTTAAGAA 2279
Qy 181 TTGTCAATCCTTACGCTTCCA 201
Db 2280 TTGTCAATCCTTACGCTTCCA 2300

RESULT 11
AAV46453 ID AAV46453 standard; cDNA; 5711 BP.
AC AAV46453;
XX
DT 18-NOV-1998 (first entry)
XX
DE Human BRCA1 om1 polymorphism #3 cDNA.
XX
KW BRCA1: om1; human: breast and ovarian cancer predisposing gene;
KW polymorphism; susceptibility; anti-oncogene; tumour suppressor;
KW chromosome 17q; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 120..5711
FT FT /*tag= a
FT FT /*product= "BRCA1 om1 protein"
FT FT 2731
FT FT variation
FT FT /*tag= b
FT FT /*note= "This polymorphic variation can be a C or T
FT FT nucleotide"
XX
PN US5750400-A.
XX
PD 12-MAY-1998.
XX
PE 12-FEB-1997; 97US-0798691.
XX
PR 12-FEB-1996; 96US-0598591.
XX PR 12-FEB-1997; 97US-0798691.
XX
PA (ONCO-) ONCORMED INC.
XX
PI Allen AC, Alvares CP, Critz BS, Murphy PD, Olson SJ;
PI Schelter DB, Zeng B;
XX
DR WPI: 1998-296774/26.
XX
XX
PT BRCA1 om1 gene coding sequences - useful for distinguishing between
PT polymorphisms and mutation(s) in the screening for disposition to
PT breast or ovarian cancer
XX
```

PS Claim 2e; Page -: 54pp; English.

XX This sequence encodes a human BRCA1 (breast and ovarian cancer  
CC predisposing gene) om1 gene in which a polymorphic variation occurs at  
CC nucleotide 2731. This sequence and other polymorphic variations of this  
CC sequence are useful for the identification of an individual who may or  
CC may not have an increased susceptibility to breast or ovarian cancer.  
CC The sequences used identify gene changes which are due to polymorphisms,  
CC rather than cancer causing mutations. BRCA1 is an anti-oncogene (tumour  
CC suppressor) which is involved in genetic inheritance of cancers, especially  
CC breast and ovarian cancer. It is found at human chromosome  
CC 17q which is known to be linked to cancer susceptibility, especially  
CC breast cancer. Cells containing a mutation in this gene lose the  
CC wild-type function of BRCA1 and are more susceptible to cancers.  
CC NOTE: This sequence does not appear in the specification but has been  
CC created from the wild type BRCA1 om1 gene represented in AAV46448.

SQ Sequence 5711 BP; 1953 A; 1099 C; 1277 G; 1381 T; 1 other;

Query Match 100.0%; Score 201; DB 19; Length 5711;  
Best Local Similarity 100.0%; Pred. No. 6.4e-97;  
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGCACAGCAGAAACCTACACTCATGAGGTAAGAACCTCGACCTGAGCCAGAG 60  
DB 2100 AGGCACAGCAGAAACCTACACTCATGAGGTAAGAACCTCGACCTGAGCCAGAG 2159

QY 61 AGTAACAGCCAAATGACAGACAGTAAGAACAGATGATGATCTTCCAGAGCTG 120  
DB 2160 AGTAACAGCCAAATGACAGACAGTAAGAACAGATGATGATCTTCCAGAGCTG 2219

QY 121 AAGTTAACAAATGACACCTGCTTCTTTACTAAGTGTTCAAATACAGTGAATTAAGAA 180  
DB 2220 AAGTTAACAAATGACACCTGCTTCTTTACTAAGTGTTCAAATACAGTGAATTAAGAA 2279

QY 181 TTTGTCAATCCTAGCCTTCCA 201  
DB 2280 TTTGTCAATCCTAGCCTTCCA 2300

RESULT 12  
AAV46454  
ID AAV46454 standard; cDNA; 5711 BP.

XX AAV46454;  
AC 18-NOV-1998 (first entry)  
DT 18-NOV-1998 (first entry)  
DE Human BRCA1 om1 polymorphism #4 cDNA.  
XX  
XX BRCA1; om1; human; breast and ovarian cancer predisposing gene;  
KM polymorphism; susceptibility; anti-oncogene; tumour suppressor;  
KW chromosome 17q; ss.  
XX  
XX Homo sapiens.  
OS  
FH Key Location/Qualifiers  
FT CDS 120..5711  
FT FT /tag= a  
FT FT /product= "BRCA1 om1 protein"  
FT FT 3232  
FT FT /tag= b  
FT FT /note= "This polymorphic variation can be an A or G  
XX nucleotide"

PM US5750400-A.  
XX  
XX 12-MAY-1998.  
PD  
XX  
XX 12-FEB-1997; 97US-0798691.  
PE  
XX 12-FEB-1996; 96US-0598591.  
PR  
PR 12-FEB-1997; 97US-0798691.

XX (ONCO-) ONCORMED INC.  
PA  
XX  
XX Allen AC, Alvares CP, Critz BS, Murphy PD, Olson SJ;  
PI Scheller DB, Zeng B;  
PI  
XX  
XX WPI: 1998-296774/26.  
DR  
XX  
PT BRCA1 om1 gene coding sequences - useful for distinguishing between  
PT polymorphisms and mutation(s) in the screening for disposition to  
PT breast or ovarian cancer  
PS  
PS Claim 2e; Page -: 54pp; English.

XX This sequence encodes a human BRCA1 (breast and ovarian cancer  
CC predisposing gene) om1 gene in which a polymorphic variation occurs at  
CC nucleotide 3232. This sequence and other polymorphic variations of this  
CC sequence are useful for the identification of an individual who may or  
CC may not have an increased susceptibility to breast or ovarian cancer.  
CC The sequences used identify gene changes which are due to polymorphisms,  
CC rather than cancer causing mutations. BRCA1 is an anti-oncogene (tumour  
CC suppressor) which is involved in genetic inheritance of cancers, especially  
CC breast and ovarian cancer. It is found at human chromosome  
CC 17q which is known to be linked to cancer susceptibility, especially  
CC breast cancer. Cells containing a mutation in this gene lose the  
CC wild-type function of BRCA1 and are more susceptible to cancers.  
CC NOTE: This sequence does not appear in the specification but has been  
CC created from the wild type BRCA1 om1 gene represented in AAV46448.

SQ Sequence 5711 BP; 1953 A; 1099 C; 1276 G; 1382 T; 1 other;

Query Match 100.0%; Score 201; DB 19; Length 5711;  
Best Local Similarity 100.0%; Pred. No. 6.4e-97;  
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGCACAGCAGAAACCTACACTCATGAGGTAAGAACCTCGACCTGAGCCAGAG 60  
DB 2100 AGGCACAGCAGAAACCTACACTCATGAGGTAAGAACCTCGACCTGAGCCAGAG 2159

QY 61 AGTAACAGCCAAATGACAGACAGTAAGAACAGATGATGATCTTCCAGAGCTG 120  
DB 2160 AGTAACAGCCAAATGACAGACAGTAAGAACAGATGATGATCTTCCAGAGCTG 2219

QY 121 AAGTTAACAAATGACACCTGCTTCTTTACTAAGTGTTCAAATACAGTGAATTAAGAA 180  
DB 2220 AAGTTAACAAATGACACCTGCTTCTTTACTAAGTGTTCAAATACAGTGAATTAAGAA 2279

QY 181 TTTGTCAATCCTAGCCTTCCA 201  
DB 2280 TTTGTCAATCCTAGCCTTCCA 2300

RESULT 13  
AAV46455  
ID AAV46455 standard; cDNA; 5711 BP.

XX AAV46455;  
AC 18-NOV-1998 (first entry)  
DT 18-NOV-1998 (first entry)  
DE Human BRCA1 om1 polymorphism #5 cDNA.  
XX  
XX  
XX BRCA1; om1; human; breast and ovarian cancer predisposing gene;  
KM polymorphism; susceptibility; anti-oncogene; tumour suppressor;  
KW chromosome 17q; ss.  
XX  
XX Homo sapiens.  
OS  
FH Key Location/Qualifiers  
FT CDS 120..5711  
FT FT /tag= a  
FT FT /product= "BRCA1 om1 protein"  
FT FT 3667  
FT FT variation

```

FT      /*tag= b
FT      /note= "This polymorphic variation can be an A or G
FT      nucleotide"
XX      US5750400-A.
XX      12-MAY-1998.
XX      12-FEB-1997; 97US-0798691.
XX      12-FEB-1996; 96US-0598591.
XX      12-FEB-1997; 97US-0798691.
XX      (ONCO-) ONCORMED INC.
XX      Allen AC, Alvares CP, Critz BS, Murphy PD, Olson SJ;
XX      Schelter DB, Zeng B;
XX      WPI; 1998-296774/26.
XX      BRCA1 om1 gene coding sequences - useful for distinguishing between
XX      PT polymorphisms and mutation(s) in the screening for disposition to
XX      PT breast or ovarian cancer
XX      Claim 2e; Page -: 54pp; English.
XX      This sequence encodes a human BRCA1 (breast and ovarian cancer
XX      CC predisposing gene) om1 gene in which a polymorphic variation occurs at
XX      CC nucleotide 3667. This sequence and other polymorphic variations of this
XX      CC sequence are useful for the identification of an individual who may or
XX      CC may not have an increased susceptibility to breast or ovarian cancer.
XX      CC The sequences used identify gene changes which are due to polymorphisms,
XX      CC rather than cancer causing mutations. BRCA1 is an anti-oncogene (tumour
XX      CC suppressor) which is involved in genetic inheritance of cancers,
XX      CC especially breast and ovarian cancer. It is found at human chromosome
XX      CC 17q which is known to be linked to cancer susceptibility, especially
XX      CC breast cancer. Cells containing a mutation in this gene lose the
XX      CC wild-type function of BRCA1 and are more susceptible to cancers.
XX      CC NOTE: This sequence does not appear in the specification but has been
XX      CC created from the wild type BRCA1 om1 gene represented in AAV46448.
XX      SQ Sequence 5711 BP; 1953 A; 1099 C; 1276 G; 1382 T; 1 other:
XX      Query Match 100.0%; Score 201; DB 19; Length 5711;
XX      Best Local Similarity 100.0%; Pred. No. 6.4e-97;
XX      Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX      QY 1 AGGCACAGCAGAACTACATCTATGGAAGTAAAGAACCTGAGCCAGAG 60
XX      DB 2100 AGGCACAGCAGAACTACATCTATGGAAGTAAAGAACCTGAGCCAGAG 2159
XX      QY 61 AGTAAACAGCCAAATGACAGACAGTAAAGACATGATATCTTCCAGAGCTG 120
XX      DB 2160 AGTAAACAGCCAAATGACAGACAGTAAAGACATGATATCTTCCAGAGCTG 2219
XX      QY 121 AAGTTACAACATGACCTGTTCTTTTACTAAGTGTCAATACCGAGTAAAGAA 180
XX      DB 2220 AAGTTACAACATGACCTGTTCTTTTACTAAGTGTCAATACCGAGTAAAGAA 2279
XX      QY 181 TTTGTCAATCTAGCCTTCCA 201
XX      DB 2280 TTTGTCAATCTAGCCTTCCA 2300
XX      RESULT 14
XX      AAV46456
XX      ID AAV46456 standard; cDNA; 5711 BP.
XX      AC AAV46456;
XX      XX
XX      DT 18-NOV-1998 (first entry)
XX      XX Human BRCA1 om1 polymorphism #6 cDNA.

```

```

XX      KW BRCA1; om1; human; breast and ovarian cancer predisposing gene;
XX      KM polymorphism; susceptibility; anti-oncogene; tumour suppressor;
XX      KW chromosome 17q; ss.
XX      OS Homo sapiens.
XX      FH Key Location/Qualifiers
XX      FT CDS 120..5711
XX      FT /tag= a
XX      FT /product= "BRCA1 om1 protein"
XX      FT 4427
XX      FT variation
XX      FT /tag= b
XX      FT /note= "This polymorphic variation can be a T or C
XX      PN US5750400-A.
XX      PD 12-MAY-1998.
XX      PF 12-FEB-1997; 97US-0798691.
XX      PR 12-FEB-1996; 96US-0598591.
XX      PR 12-FEB-1997; 97US-0798691.
XX      PA (ONCO-) ONCORMED INC.
XX      PI Allen AC, Alvares CP, Critz BS, Murphy PD, Olson SJ;
XX      PI Schelter DB, Zeng B;
XX      PT WPI; 1998-296774/26.
XX      PT BRCA1 om1 gene coding sequences - useful for distinguishing between
XX      PT polymorphisms and mutation(s) in the screening for disposition to
XX      PT breast or ovarian cancer
XX      Claim 2e; Page -: 54pp; English.
XX      This sequence encodes a human BRCA1 (breast and ovarian cancer
XX      CC predisposing gene) om1 gene in which a polymorphic variation occurs at
XX      CC nucleotide 4427. This sequence and other polymorphic variations of this
XX      CC sequence are useful for the identification of an individual who may or
XX      CC may not have an increased susceptibility to breast or ovarian cancer.
XX      CC The sequences used identify gene changes which are due to polymorphisms,
XX      CC rather than cancer causing mutations. BRCA1 is an anti-oncogene (tumour
XX      CC suppressor) which is involved in genetic inheritance of cancers,
XX      CC especially breast and ovarian cancer. It is found at human chromosome
XX      CC 17q which is known to be linked to cancer susceptibility, especially
XX      CC breast cancer. Cells containing a mutation in this gene lose the
XX      CC wild-type function of BRCA1 and are more susceptible to cancers.
XX      CC NOTE: This sequence does not appear in the specification but has been
XX      CC created from the wild type BRCA1 om1 gene represented in AAV46448.
XX      SQ Sequence 5711 BP; 1953 A; 1098 C; 1277 G; 1382 T; 1 other:
XX      Query Match 100.0%; Score 201; DB 19; Length 5711;
XX      Best Local Similarity 100.0%; Pred. No. 6.4e-97;
XX      Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX      QY 1 AGGCACAGCAGAACTACATCTATGGAAGTAAAGAACCTGAGCCAGAG 60
XX      DB 2100 AGGCACAGCAGAACTACATCTATGGAAGTAAAGAACCTGAGCCAGAG 2159
XX      QY 61 AGTAAACAGCCAAATGACAGACAGTAAAGACATGATATCTTCCAGAGCTG 120
XX      DB 2160 AGTAAACAGCCAAATGACAGACAGTAAAGACATGATATCTTCCAGAGCTG 2219
XX      QY 121 AAGTTACAACATGACCTGTTCTTTTACTAAGTGTCAATACCGAGTAAAGAA 180
XX      DB 2220 AAGTTACAACATGACCTGTTCTTTTACTAAGTGTCAATACCGAGTAAAGAA 2279
XX      QY 181 TTTGTCAATCTAGCCTTCCA 201
XX      DB 2280 TTTGTCAATCTAGCCTTCCA 2300

```

Db 2280 TTTGTCAATCCTAGCCTTCCA 2300

RESULT 15  
ADVANCE7

ID	AAV46457 standard; CDNA; 5711 BP.

AC AAV46457;

DT 18-NOV-1998 (first entry)  
XX

Human BRCA1 om1 polymorphism #7 cDNA

KW BRCA1; oml; human; breast and ovarian cancer predisposing gene  
KW polymorphism; susceptibility; anti-oncogene; tumour suppressor;  
KW chromosome 17q; ss.

OS Homo sapiens.

FH	Key	Location/Qualifiers
FT	CDS	120..5711

120..5711

FT /product= "BRCA1 omil protein"

FT variation

FT /note= "This polymorphic variation can be an A or G  
FT nucleotide"

PN US5750400-A.

PD 12-MAY-1998.

PF 12-FEB-1997; 97US-0798691.

PR 12-FEB-1996; 96US-0598591.

PR 12-FEB-1997; 97US-0798691.  
VY

PA (ONCO-) ONCORMED INC.  
XX

PI Allen AC, Alvares CP, Critz BS, Murphy PD, Olson SJ;  
PI Scholten DB, Zeng B.

PI Schelter DB, Zeng B;  
xx

WPI; 1.998-296774/26.

PT BRCA1 omi gene coding sequences - useful for distinguishing between  
PT polymorphisms and mutation(s) in the screening for disposition to  
PT breast or ovarian cancer

PS Claim 2e; Page -: 54pp; English.  
xy

This sequence encodes a human BRCA1 (breast and ovarian cancer predisposing gene) om1 gene in which a polymorphic variation occurs at nucleotide 4956. This sequence and other polymorphic variations of this sequence are useful for the identification of an individual who may or may not have an increased susceptibility to breast or ovarian cancer. The sequences used identify gene changes which are due to polymorphisms rather than cancer causing mutations. BRCA1 is an anti-oncogene (tumour suppressor) which is involved in genetic inheritance of cancers, especially breast and ovarian cancer. It is found at human chromosome 17q which is known to be linked to cancer susceptibility, especially breast cancer. Cells containing a mutation in this gene lose the wild-type function of BRCA1 and are more susceptible to cancers. NOTE: This sequence does not appear in the specification but has been created from the wild type BRCA1 om1 gene represented in AA964448.

SQ Sequence 5711..BP; 1953 A; 1099 C; 1276 G; 1382 T; 1 other;

Query Match	100.0%	Score 201;	DB 19;	Length 5711;
-------------	--------	------------	--------	--------------

Best Local Similarity 100.08; Pred. No. 6.4e-97;

Matches	201;	Conservative	0;	Mismatches	0;	Indels	0;	. Gaps	0;
---------	------	--------------	----	------------	----	--------	----	--------	----

```

Db      2100  AGGCACAGCAAAAACCTCACTCACTGTGGAAGSTTAAGAAACCTGCAACTGGAGCCAAAG  2159
QY      61    AGTAACAGCCAAATGAAACAGACAAAGTAAAGACATGACAGTACTCTTCCAGAGCTG  120
        |||||
Db      2160  AGTAAACAAGCCAAATGAACACAGCAAGTAAAGACATGACAGTGAATCTTCCAGAGCTG  2219
QY      121  AAGTTAACAAATGACACCTGCTTCTTTACTAAGTTCAAAATACCAGAGCACTTAAAGAA  180
        |||||
Db      2220  AAGTTAACAAATGACACCTGCTTCTTTACTAAGTTCAAAATACCAGAGCACTTAAAGAA  2279
QY      181  TTTGTCAATCCTAGCCTTCCA  201
        |||||
Db      2280  TTTGTCAATCCTAGCCTTCCA  2300
        .

Search completed: June 27, 2003, 05:28:56
Job time : 248 secs

```

61 AGTAACAAGCCAAATGAACAGACAAGTAAAGACATGACAGTGATACTTTCCAGAGCTG 120

Db 2160 AGTACACAGCCAAATGACACAGACCAAGTAAAAAGACATGACAGTGATACTTCCAGAGCTG 2219

121 AAGTTAACAATGCACCTGGTCTTTACTAAGTGTCAATACCAGTGAACCTAAAGAA 180

Db 2220 AAGTTAACAAATGCACCCTGGTTCCTTTTACTAAGTGTTCAAATACCAGTGAACCTTAAGAA 2279

181 TTGTCATCCTAGCCTCCA 201

Db 2280 TTGTCAATCCTAGCCTCCA 2300

```
Search completed: June 27, 2003, 05:28:56
Job time : 248 secs
```

Job time : 248 secs



Fri Jun 27 10:47:45 2003

us-09-734-672-3\_copy\_2100\_2300.01.mpb

C to T at 2201

Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 27, 2003, 01:02:23 ; Search time 149 Seconds  
(Without alignments)  
2002.511 Million cell updates/sec

Title: US-09-734-672-3\_COPY\_2100\_2300  
Perfect score: 201  
Sequence: 1 AGGCACAGCAGAACTACA.....TTGTCAATCCTACCTTCCA 201

Scoring table:  
Gapop 60.0 , Gapext 60.0

Searched: 1055720 seqs, 742224136 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2111440

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Published\_Applications\_NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	201	100.0	5711	9	US-09-734-672-1	Sequence 1, Appl1
2	201	100.0	5711	9	US-09-734-672-5	Sequence 5, Appl1
3	201	100.0	5711	9	US-09-982-828-1	Sequence 1, Appl1
4	201	100.0	5711	9	US-09-982-828-3	Sequence 3, Appl1
5	201	100.0	5711	9	US-10-022-819-1	Sequence 1, Appl1
6	150	74.6	5711	9	US-09-734-672-3	Sequence 3, Appl1
7	150	74.6	5711	9	US-09-982-828-5	Sequence 5, Appl1
8	84	41.8	121	9	US-09-818-875-765	Sequence 765, App
9	84	41.8	121	9	US-09-818-875-766	Sequence 766, App
10	70	34.8	121	9	US-09-818-875-761	Sequence 761, App
11	70	34.8	121	9	US-09-818-875-762	Sequence 762, App
12	31	15.4	31	10	US-09-838-497-8	Sequence 8, Appl1
13	24	11.4	24	9	US-10-259-479-3	Sequence 32, Appl1
14	23	11.4	23	9	US-09-734-672-32	Sequence 32, Appl1
15	23	11.4	23	9	US-09-982-828-32	Sequence 28, Appl1
16	23	11.4	23	9	US-10-022-819-28	Sequence 96, Appl1
17	23	11.4	23	9	US-09-923-327-96	Sequence 1, Appl1
18	20	10.0	24	9	US-10-259-479-1	Sequence 2378, Appl1
19	19	9.5	282	10	US-09-294-093B-2378	Sequence 2378, Appl1

20	18	9.0	293	10	US-09-867-701-339	Sequence 339, App
21	18	9.0	309	10	US-09-867-701-410	Sequence 410, App
22	18	9.0	505	10	US-09-864-761-13432	Sequence 13432, A
23	18	9.0	1763	10	US-09-917-800A-1460	Sequence 1460, Ap
24	17	8.5	17	9	US-09-818-875-767	Sequence 767, App
25	17	8.5	17	9	US-09-818-875-768	Sequence 768, App
26	17	8.5	17	9	US-09-923-327-170	Sequence 170, App
27	17	8.5	17	9	US-09-923-327-189	Sequence 189, App
28	17	8.5	237	9	US-10-054-387-74	Sequence 74, Appl
29	17	8.5	284	10	US-09-878-574-8061	Sequence 8061, Ap
30	17	8.5	648	9	US-09-991-936-1599	Sequence 1599, Ap
31	17	8.5	4842	9	US-10-174-590-289	Sequence 289, App
32	17	8.5	4842	9	US-10-176-758-289	Sequence 289, App
33	17	8.5	4842	9	US-10-175-737-289	Sequence 289, App
34	17	8.5	4842	9	US-10-173-706-289	Sequence 289, App
35	17	8.5	4842	9	US-10-175-738-289	Sequence 289, App
36	17	8.5	4842	9	US-10-175-738-289	Sequence 289, App
37	17	8.5	4842	9	US-10-176-482-289	Sequence 289, App
38	17	8.5	4842	9	US-10-176-482-289	Sequence 289, App
39	17	8.5	4842	9	US-10-176-757-289	Sequence 289, App
40	17	8.5	4842	9	US-10-180-552-289	Sequence 289, App
41	17	8.5	4842	9	US-10-180-557-289	Sequence 289, App
42	17	8.5	4842	9	US-10-173-700-289	Sequence 289, App
43	17	8.5	4842	9	US-10-174-572-289	Sequence 289, App
44	17	8.5	4842	9	US-10-174-579-289	Sequence 289, App
45	17	8.5	4842	9	US-10-174-582-289	Sequence 289, App

# ALIGNMENTS

RESULT 1  
US-09-734-672-1  
Sequence 1, Application US/09734672  
Publication No. US20020183268A1  
GENERAL INFORMATION:  
APPLICANT: Murphy, Patricia D.  
Allen, Antonette C.  
Critz, Brenda S.  
Olson, Sheri J.  
Schelter, Denise B.  
Zeng, Bin  
TITLE OF INVENTION: Coding Sequences of the Human  
BRCA1 Gene  
NUMBER OF SEQUENCES: 72  
CORRESPONDENCE ADDRESS:  
ADDRESS: Morgan Lewis & Bockius LLP  
STREET: 1111 Pennsylvania Ave., N.W.  
City: Washington  
STATE: District of Columbia  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/734,672  
FILING DATE: 03-Dec-2000  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/966,436  
FILING DATE: 07-No. US20020183268A1-97  
APPLICATION NUMBER: US 08/598,591  
FILING DATE: 12-Feb-96  
ATTORNEY/AGENT INFORMATION:  
NAME: Michael S. Tuscan  
REGISTRATION NUMBER: 43,210  
REFERENCE/DOCKET NUMBER: 44921-5055-02-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-739-3000  
TELEFAX: 202-739-3001

INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5711 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: No. US20020183268A1 Relevant  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
STRAIN: BRCAl  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: 17  
MAP POSITION: 17q21  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-734-672-1

Query Match 100.0%; Score 201; DB 9; Length 5711;  
Best Local Similarity 100.0%; Pred. No. 8.9e-100;  
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGCACAGCAGAACTACACTGAGGTAAAGCACTGCACTGGAGCCAGAG 60  
DB 2100 AGGCACAGCAGAACTACACTGAGGTAAAGCACTGCACTGGAGCCAGAG 2159  
QY 61 AGTAACAAGCCAAATGACAGACAGTAAGACATGACATGATCTTCCAGAGCTG 120  
DB 2160 AGTAACAAGCCAAATGACAGACAGTAAGACATGACATGATCTTCCAGAGCTG 2219  
QY 121 AAGTTAACAAATGACACCTGTTCTTTACTAAGTGTCAATACGAGTAAAGAA 180  
DB 2220 AAGTTAACAAATGACACCTGTTCTTTACTAAGTGTCAATACGAGTAAAGAA 2279  
QY 181 TTGTCAATCCTAGCCTTCCA 201  
DB 2280 TTGTCAATCCTAGCCTTCCA 2300

RESULT 2  
US-09-734-672-5  
Sequence 5, Application US/09734672  
Publication No. US20020183268A1  
GENERAL INFORMATION:  
APPLICANT: Murphy, Patricia D.  
Allen, Antonette C.  
Alvares, Christopher P.  
Critz, Brenda S.  
Olson, Sheri J.  
Schelter, Denise B.  
Zeng, Bin  
TITLE OF INVENTION: Coding Sequences of the Human  
BRCAl Gene  
NUMBER OF SEQUENCES: 72  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morgan Lewis & Bockius LLP  
STREET: 1111 Pennsylvania Ave., N.W.  
CITY: Washington  
STATE: District of Columbia  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/734,672  
FILING DATE: 03-Dec-2000  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/966,436  
FILING DATE: 07-No. US20020183268A1-97  
APPLICATION NUMBER: US 08/596,591  
FILING DATE: 12-Feb-96  
ATTORNEY/AGENT INFORMATION:

NAME: Michael S. Tuscan  
REGISTRATION NUMBER: 43,210  
REFERENCE/DOCKET NUMBER: 44921-5055-02-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-739-3000  
TELEFAX: 202-739-3001  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5711 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: No. US20020183268A1 Relevant  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
STRAIN: BRCAl  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: 17  
MAP POSITION: 17q21  
SEQUENCE DESCRIPTION: SEQ ID NO: 5:  
US-09-734-672-5

Query Match 100.0%; Score 201; DB 9; Length 5711;  
Best Local Similarity 100.0%; Pred. No. 8.9e-100;  
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGCACAGCAGAACTACACTGAGGTAAAGCACTGCACTGGAGCCAGAG 60  
DB 2100 AGGCACAGCAGAACTACACTGAGGTAAAGCACTGCACTGGAGCCAGAG 2159  
QY 61 AGTAACAAGCCAAATGACAGACAGTAAGACATGACATGATCTTCCAGAGCTG 120  
DB 2160 AGTAACAAGCCAAATGACAGACAGTAAGACATGACATGATCTTCCAGAGCTG 2219  
QY 121 AAGTTAACAAATGACACCTGTTCTTTACTAAGTGTCAATACGAGTAAAGAA 180  
DB 2220 AAGTTAACAAATGACACCTGTTCTTTACTAAGTGTCAATACGAGTAAAGAA 2279  
QY 181 TTGTCAATCCTAGCCTTCCA 201  
DB 2280 TTGTCAATCCTAGCCTTCCA 2300

RESULT 3  
US-09-982-828-1  
Sequence 1, Application US/09982828  
Publication No. US20030022184A1  
GENERAL INFORMATION:  
APPLICANT: Murphy, Patricia D.  
Allen, Antonette C.  
Alvares, Christopher P.  
Critz, Brenda S.  
Olson, Sheri J.  
Thurber, Denise  
Zeng, Bin  
TITLE OF INVENTION: Coding Sequences of the Human  
BRCAl Gene  
NUMBER OF SEQUENCES: 72  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morgan Lewis & Bockius LLP  
STREET: 1111 Pennsylvania Avenue N. W.  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/982,828  
FILING DATE: 22-Oct-2001

ZIP: 20004  
COMPUTER.READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/982,828  
FILING DATE: 22-Oct-2001  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/074,453  
FILING DATE: 1998-05-06  
APPLICATION NUMBER: US 08/798,691  
FILING DATE: 1997-02-12  
APPLICATION NUMBER: US 08/598,591  
FILING DATE: 1996-02-12  
ATTORNEY/AGENT INFORMATION:  
NAME: Michael S. Tuscan  
REGISTRATION NUMBER: 43,210  
REFERENCE/DOCKET NUMBER: 44921-5053-01-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-739-3000  
TELEFAX: 202-739-3001  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 571 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear

```

: ORIGINAL SOURCE:
: ORGANISM: Homo sapiens
: STRAIN: BRCA1 (om12)
: POSITION IN GENOME:
: CHROMOSOME/SEGMENT: 17
: MAP POSITION: 17q21
: SEQUENCE DESCRIPTION: SEQ ID NO: 3
: US-09-982-828-3

Query Match      100.0%; Score 201; DB 9; Length 5711;
Best Local Similarity 100.0%; Pred. No. 8; 9e-100;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AGGCACAGCAGAAACCTACACACTCATGTGAAGGTAAAGAACTCGCACTGGAGCCAGAAG 60
      |||||
Db      2100 AGGCACAGCAGAAACCTACACACTCATGTGAAGGTAAAGAACTCGCACTGGAGCCAGAAG 2159

QY      61 AGTACACAGCCAAATGTAACAGACAGTAAGAAAGACATGACAGTGTACTTTCAGAGACTG 120
      |||||

```

OY	121	AACTTAAACAATGCACCGTGGTCTTTACTAAGATGCTCAATGCCACTGAACCTAAAGAA	180
Db	2220	AAGTTAACAAATGACCGCTGGTCTTTACTAAGTGTCAATACCACTGAACCTAAAGAA	2279
OY	181	TTTGTCATCTAGCCCTTCCA	201
Db	2280	TTTGTCATCTAGCCCTTCCA	2300
RESULT 5			
	US-10-022-819-1		
	Sequence 1, Application US/10022819		
	Publication No. US2003027166A1		
	GENERAL INFORMATION:		
	APPLICANT: ALLEN, Antonette C. P.		
	OLSEN, Sheri J.		
	LAWRENCE, Tammy S.		
	ANGELLY, Tracy S.		
	RABIN, Mark B.		
	TITLE OF INVENTION: CODING SEQUENCE HAPLOTYPE OF THE HUMAN		
	BRCA1 GENE		

NUMBER OF SEQUENCES: 67  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morgan Lewis & Bockius LLP  
STREET: 1111 Pennsylvania Avenue  
CITY: Washington DC  
STATE: District of Columbia  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/022,819  
FILING DATE: 22-Apr-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/074,452  
FILING DATE: 1998-05-06  
ATTORNEY/AGENT INFORMATION:  
NAME: <Unknown>  
REGISTRATION NUMBER: <Unknown>  
REFERENCE/DOCKET NUMBER: 044921-5049-01-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-739-3000  
TELEFAX: 202-739-3001  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5711 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: HOMO SAPIENS  
STRAIN: BRCAL  
HAPLOTYPE: OM14  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: 17  
MAP POSITION: 17q21  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-10-022-819-1  
Query Match 100.0%; Score 201; DB 9; Length 5711;  
Best Local Similarity 100.0%; Pred. No. 8.9e-100;  
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGCACAGCAGAACTACACTGAGTGAAGTAAGAACCTGCACTGGAGCCAGAG 60  
|||||  
DB 2100 AGGCACAGCAGAACTACACTGAGTGAAGTAAGAACCTGCACTGGAGCCAGAG 2159  
QY 61 ACTAACAGCACAATGACAGACAGCAAGTAAGAACATGACAGTATCTTCCAGAGCTG 120  
|||||  
DB 2160 AGTAACAAGCACAATGACAGACAGCAAGTAAGAACATGACAGTATCTTCCAGAGCTG 2219  
QY 121 AAGTTAACAAATGACACCTGTTCTTTACTAAGTGTTCACAATACCAAGTGAAGAA 180  
|||||  
DB 2220 AAGTTAACAAATGACACCTGTTCTTTACTAAGTGTTCACAATACCAAGTGAAGAA 2279  
QY 181 TTTGTCAATCTAGCCTTCCA 201  
|||||  
DB 2280 TTTGTCAATCTAGCCTTCCA 2300

RESULT 6  
US-09-734-672-3  
Sequence 3, Application US/09734672  
Publication No. US20020183268A1  
GENERAL INFORMATION:

APPLICANT: Murphy, Patricia D.  
Allen, Antonette C.  
Alvares, Christopher P.  
Critz, Brenda S.  
Olson, Sheri J.  
Scheller, Denise B.  
Zeng, Bin  
TITLE OF INVENTION: Coding Sequences of the Human  
BRCAL Gene  
NUMBER OF SEQUENCES: 72  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morgan Lewis & Bockius LLP  
STREET: 1111 Pennsylvania Ave., N.W.  
CITY: Washington  
STATE: District of Columbia  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/734,672  
FILING DATE: 03-Dec-2000  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/966,436  
FILING DATE: 07-No. US20020183268A1-97  
APPLICATION NUMBER: US 08/598,591  
FILING DATE: 12-Feb-96  
ATTORNEY/AGENT INFORMATION:  
NAME: Michael S. Tuscan  
REGISTRATION NUMBER: 43,210  
REFERENCE/DOCKET NUMBER: 44921-5055-02-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-739-3000  
TELEFAX: 202-739-3001  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5711 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: No. US20020183268A1 Relevant  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
STRAIN: BRCAL  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: 17  
MAP POSITION: 17q21  
SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-09-734-672-3  
Query Match 74.6%; Score 150; DB 9; Length 5711;  
Best Local Similarity 99.5%; Pred. No. 6.4e-72;  
Matches 200; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGGCACAGCAGAACTACACTGAGTGAAGTAAGAACCTGCAACTGGAGCCAGAG 60  
|||||  
DB 2100 AGGCACAGCAGAACTACACTGAGTGAAGTAAGAACCTGCAACTGGAGCCAGAG 2159  
QY 61 ACTAACAGCACAATGACAGACAGCAAGTAAGAACATGACAGTATCTTCCAGAGCTG 120  
|||||  
DB 2160 AGTAACAAGCACAATGACAGACAGCAAGTAAGAACATGACAGTATCTTCCAGAGCTG 2219  
QY 121 AAGTTAACAAATGACACCTGTTCTTTACTAAGTGTTCACAATACCAAGTGAAGAA 180  
|||||  
DB 2220 AAGTTAACAAATGACACCTGTTCTTTACTAAGTGTTCACAATACCAAGTGAAGAA 2279  
QY 181 TTTGTCAATCTAGCCTTCCA 201  
|||||  
DB 2280 TTTGTCAATCTAGCCTTCCA 2300

RESULT 7  
US-09-828-828-5  
: Sequence 5, Application US/09982828  
: Publication No. US20030022184A1  
GENERAL INFORMATION:  
APPLICANT: Murphy, Patricia D.  
Allen, Antonette C.  
Alvares, Christopher P.  
Critz, Brenda S.  
Olson, Sheri J.  
Thurber, Denise  
Zeng, Bin  
TITLE OF INVENTION: Coding Sequences of the Human  
BRCA1 Gene  
NUMBER OF SEQUENCES: 72  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morgan Lewis & Bockius LLP  
STREET: 1111 Pennsylvania Avenue N. W.  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/982,828  
FILING DATE: 22-Oct-2001  
CLASSIFICATION: <Unknown>  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 09/074,453  
FILING DATE: 1998-05-06  
APPLICATION NUMBER: US 08/798,691  
FILING DATE: 1997-02-12  
APPLICATION NUMBER: US 08/598,591  
FILING DATE: 1996-02-12  
ATTORNEY/AGENT INFORMATION:  
NAME: Michael S. Tuscan  
REGISTRATION NUMBER: 43,210  
REFERENCE/DOCKET NUMBER: 44921-5053-01-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-739-3000  
TELEFAX: 202-739-3001  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5711 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
STRAIN: BRCA1 (om13)  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: 17  
MAP POSITION: 17q21  
SEQUENCE DESCRIPTION: SEQ ID NO: 5:  
US-09-828-828-5  
Query Match 74.6% Score 150; DB 9; Length 5711;  
Best Local Similarity 99.5%; Pred. No. 6.4e-72;  
Matches 200; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGGCACAGCAGAACTACATGAGAGTAAAGACCTCACTGGAGCCAGAAG 60  
|||||  
DB 2100 AGGCACAGCAGAACTACATGAGAGTAAAGACCTCACTGGAGCCAGAAG 2159  
QY 61 AGTAAACAGCAGCAATGATGAGAGCAAGTAAAGACATGACGTATCTTTCCAGAGCTG 120  
|||||  
DB 2160 AGTAAACAGCAGCAATGATGAGAGCAAGTAAAGACATGACGTATCTTTCCAGAGCTG 2219

QY 121 AAGTTACAAATGACACCTGGTCTTTTACTAGTGTCAATACCACTGAACCTTAAGAA 180  
|||||  
DB 2220 AAGTTACAAATGACACCTGGTCTTTTACTAGTGTCAATACCACTGAACCTTAAGAA 2279  
QY 181 TTTGTCAATCCTAGACCTTCCA 201  
|||||  
DB 2280 TTTGTCAATCCTAGACCTTCCA 2300

RESULT 8  
US-09-818-875-765  
: Sequence 765, Application US/09818875  
: Publication No. US20030051270A1  
GENERAL INFORMATION:  
APPLICANT: Kmiec, Eric B.  
Applicant: Gamper, Howard B.  
Applicant: Rice, Michael C.  
TITLE OF INVENTION: Targeted Chromosomal Genomic Alterations with Modified Single  
FILE REFERENCE: Napro-4  
CURRENT APPLICATION NUMBER: US/09/818,875  
CURRENT FILING DATE: 2001-03-27  
PRIOR APPLICATION NUMBER: US 60/192,176  
PRIOR FILING DATE: 2000-03-27  
PRIOR APPLICATION NUMBER: US 60/192,179  
PRIOR FILING DATE: 2000-03-27  
PRIOR APPLICATION NUMBER: US 60/208,538  
PRIOR FILING DATE: 2000-06-01  
PRIOR APPLICATION NUMBER: US 60/244,989  
NUMBER OF SEQ ID NOS: 4385  
SOFTWARE: Friedman macro Napro4  
SEQ ID NO 765  
LENGTH: 121  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-818-875-765  
Query Match 41.8% Score 84; DB 9; Length 121;  
Best Local Similarity 100.0%; Pred. No. 5.9e-36;  
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 118 CTGAGTTACAAATGACACCTGGTCTTTTACTAGTGTCAATACCACTGAACCTTAA 177  
|||||  
DB 1 CTGAGTTACAAATGACACCTGGTCTTTTACTAGTGTCAATACCACTGAACCTTAA 60  
QY 178 GAATTTGTCAATCCTAGACCTTCCA 201  
|||||  
DB 61 GAATTTGTCAATCCTAGACCTTCCA 84

RESULT 9  
US-09-818-875-766/c  
: Sequence 766, Application US/09818875  
: Publication No. US20030051270A1  
GENERAL INFORMATION:  
APPLICANT: Kmiec, Eric B.  
Applicant: Gamper, Howard B.  
Applicant: Rice, Michael C.  
TITLE OF INVENTION: Targeted Chromosomal Genomic Alterations with Modified Single  
FILE REFERENCE: Napro-4  
CURRENT APPLICATION NUMBER: US/09/818,875  
CURRENT FILING DATE: 2001-03-27  
PRIOR APPLICATION NUMBER: US 60/192,176  
PRIOR FILING DATE: 2000-03-27  
PRIOR APPLICATION NUMBER: US 60/192,179  
PRIOR FILING DATE: 2000-03-27  
PRIOR APPLICATION NUMBER: US 60/208,538  
PRIOR FILING DATE: 2000-06-01  
PRIOR APPLICATION NUMBER: US 60/244,989  
PRIOR FILING DATE: 2000-10-30

NUMBER OF SEQ ID NOS: 4385  
 SOFTWARE: Friedmann macro Napro4  
 SEQ ID NO 766  
 LENGTH: 121  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 US-09-818-875-766

Query Match  
 Best Local Similarity 41.8%; Score 84; DB 9; Length 121;  
 Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 118 CTGAAGTAAACAATGACCTGTTCTTTTACTAAGTGTCAATACCACTGAAGTAA 177  
 |||||||  
 Db 121 CTGAAGTAAACAATGACCTGTTCTTTTACTAAGTGTCAATACCACTGAAGTAA 62

QY 178 GAATTTGTCATCTAGCTTCCA 201  
 |||||||  
 Db 61 GAATTTGTCATCTAGCTTCCA 38

RESULT 10  
 US-09-818-875-761  
 Sequence 761, Application US/09818875  
 Publication No. US20030051270A1  
 GENERAL INFORMATION:

APPLICANT: Kmiec, Eric B.  
 APPLICANT: Gamper, Howard B.  
 APPLICANT: Rice, Michael C.  
 TITLE OF INVENTION: Targeted Chromosomal Genomic Alterations with Modified Single  
 FILE REFERENCE: Napro-4  
 CURRENT APPLICATION NUMBER: US/09/818,875  
 CURRENT FILING DATE: 2001-03-27  
 PRIOR APPLICATION NUMBER: US 60/192,176  
 PRIOR FILING DATE: 2000-03-27  
 PRIOR APPLICATION NUMBER: US 60/192,179  
 PRIOR FILING DATE: 2000-03-27  
 PRIOR APPLICATION NUMBER: US 60/208,538  
 PRIOR FILING DATE: 2000-06-01  
 PRIOR APPLICATION NUMBER: US 60/244,989  
 PRIOR FILING DATE: 2000-10-30  
 NUMBER OF SEQ ID NOS: 4385  
 SOFTWARE: Friedmann macro Napro4  
 SEQ ID NO 761  
 LENGTH: 121  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 US-09-818-875-761

Query Match  
 Best Local Similarity 34.8%; Score 70; DB 9; Length 121;  
 Matches 120; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 37 GAACCTGCACTGGAGCCAGAGTAACAAGCCCAATGACAGACAGTAAGACAT 96  
 |||||||  
 Db 1 GAACCTGCACTGGAGCCAGAGTAACAAGCCCAATGACAGACAGTAAGACAT 60

QY 97 GACAGTATCTTCCAGAGCTGAAGTTAACAATGACCCGGTCTTTCTACTAGTGT 156  
 |||||||  
 Db 61 GACAGTATCTTCCAGAGCTGAAGTTAACAATGACCCGGTCTTTCTACTAGTGT 120

QY 157 T 157  
 Db 121 T 121

RESULT 11  
 US-09-818-875-762/C  
 Sequence 762, Application US/09818875  
 Publication No. US20030051270A1  
 GENERAL INFORMATION:  
 APPLICANT: Kmiec, Eric B.

APPLICANT: Gamper, Howard B.  
 APPLICANT: Rice, Michael C.  
 TITLE OF INVENTION: Targeted Chromosomal Genomic Alterations with Modified Single  
 FILE REFERENCE: Napro-4  
 CURRENT APPLICATION NUMBER: US/09/818,875  
 CURRENT FILING DATE: 2001-03-27  
 PRIOR APPLICATION NUMBER: US 60/192,176  
 PRIOR FILING DATE: 2000-03-27  
 PRIOR APPLICATION NUMBER: US 60/192,179  
 PRIOR FILING DATE: 2000-03-27  
 PRIOR APPLICATION NUMBER: US 60/208,538  
 PRIOR FILING DATE: 2000-06-01  
 PRIOR APPLICATION NUMBER: US 60/244,989  
 PRIOR FILING DATE: 2000-10-30  
 NUMBER OF SEQ ID NOS: 4385  
 SOFTWARE: Friedmann macro Napro4  
 SEQ ID NO 762  
 LENGTH: 121  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 US-09-818-875-762

Query Match  
 Best Local Similarity 34.8%; Score 70; DB 9; Length 121;  
 Matches 120; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 37 GAACCTGCACTGGAGCCAGAGTAACAAGCCCAATGACAGACAGTAAGACAT 96  
 |||||||  
 Db 121 GAACCTGCACTGGAGCCAGAGTAACAAGCCCAATGACAGACAGTAAGACAT 62

QY 97 GACAGTATCTTCCAGAGCTGAAGTTAACAATGACCCGGTCTTTCTACTAGTGT 156  
 |||||||  
 Db 61 GACAGTATCTTCCAGAGCTGAAGTTAACAATGACCCGGTCTTTCTACTAGTGT 2

QY 157 T 157  
 Db 1 T 1

RESULT 12  
 US-09-838-497-8  
 Sequence 8, Application US/09838497  
 Patent No. US20020061523A1  
 GENERAL INFORMATION:  
 APPLICANT: Wu, Tian  
 APPLICANT: Schwartz, David C.  
 TITLE OF INVENTION: Method for Analyzing Nucleic Acid Reactions  
 FILE REFERENCE: 09820148  
 CURRENT APPLICATION NUMBER: US/09/838,497  
 CURRENT FILING DATE: 2001-04-19  
 PRIOR APPLICATION NUMBER: US 09/175,824  
 PRIOR FILING DATE: 1998-10-20  
 NUMBER OF SEQ ID NOS: 9  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO 8  
 LENGTH: 31  
 TYPE: DNA  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Description of Artificial Sequence: primer  
 US-09-838-497-8

Query Match  
 Best Local Similarity 15.4%; Score 31; DB 10; Length 31;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 CCAGTGAAGTAAAGATTTGTCAATCTAG 194  
 |||||||  
 Db 1 CCAGTGAAGTAAAGATTTGTCAATCTAG 31

RESULT 13

US-10-259-479-3  
; Sequence 3, Application US/10259479  
; Publication No. US20030082616A1  
; GENERAL INFORMATION:  
; APPLICANT: HITACH, LTD.  
; APPLICANT: National Cancer Center  
; TITLE OF INVENTION: Genetic Screening Method and Genetic Screening Apparatus  
; FILE REFERENCE: PH-1080  
; CURRENT APPLICATION NUMBER: US/10/259,479  
; CURRENT FILING DATE: 2002-09-30  
; PRIOR APPLICATION NUMBER: JP 294257/1999  
; PRIOR FILING DATE: 1999-10-15  
; NUMBER OF SEQ ID NOS: 6  
; SEQ ID NO 3  
; LENGTH: 24  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: DNA primer used for PCR.  
US-10-259-479-3

Query Match 11.9%; Score 24; DB 9; Length 24;  
Best Local Similarity 100.0%; Pred. No. 0.0032;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 GCACGTGAGCCAGAGAGATAAC 66  
Db 1 GCACGTGAGCCAGAGAGATAAC 24

## RESULT 14

US-09-734-672-32/C  
; Sequence 32, Application US/09734672  
; Publication No. US20020183268A1

## GENERAL INFORMATION:

APPLICANT: Murphy, Patricia D.  
Allen, Antonette C.  
Alvares, Christopher P.  
Critz, Brenda S.  
Olson, Sheri J.  
Schelter, Denise B.  
Zeng, Bin

TITLE OF INVENTION: Coding Sequences of the Human  
BRCA1 Gene

NUMBER OF SEQUENCES: 72  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Morgan Lewis & Bockius LLP  
STREET: 1111 Pennsylvania Ave., N.W.  
CITY: Washington  
STATE: District of Columbia  
COUNTRY: USA  
ZIP: 20004

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/734,672  
FILING DATE: 03-Dec-2000

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/966,436  
FILING DATE: 07-No. US20020183268A1-97  
APPLICATION NUMBER: US 08/598,591  
FILING DATE: 12-Feb-96

## ATTORNEY/AGENT INFORMATION:

NAME: Michael S. Tuscan  
REGISTRATION NUMBER: 43,210  
REFERENCE/DOCKET NUMBER: 44921-5055-02-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-739-3000  
TELEFAX: 202-739-3001  
INFORMATION FOR SEQ ID NO: 32:

SEQUENCE CHARACTERISTICS:  
LENGTH: 23 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: No. US20020183268A1 Relevant  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
STRAIN: 11ER primer  
SEQUENCE DESCRIPTION: SEQ ID NO: 32:  
US-09-734-672-32

Query Match 11.4%; Score 23; DB 9; Length 23;  
Best Local Similarity 100.0%; Pred. No. 0.011;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 TGTTCATATACGATGACTTAA 176  
Db 23 TGTTCATATACGATGACTTAA 1

## RESULT 15

US-09-982-828-32/C  
; Sequence 32, Application US/09982828  
; Publication No. US20030022184A1

## GENERAL INFORMATION:

APPLICANT: Murphy, Patricia D.  
Allen, Antonette C.  
Alvares, Christopher P.  
Critz, Brenda S.  
Olson, Sheri J.  
Thurber, Denise  
Zeng, Bin

TITLE OF INVENTION: Coding Sequences of the Human  
BRCA1 Gene

NUMBER OF SEQUENCES: 72  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Morgan Lewis & Bockius LLP  
STREET: 1111 Pennsylvania Avenue N. W.  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20004

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/982,828  
FILING DATE: 22-Oct-2001  
CLASSIFICATION: <Unknown>

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 09/074,453  
FILING DATE: 1998-05-06  
APPLICATION NUMBER: US 08/798,691  
FILING DATE: 1997-02-12  
APPLICATION NUMBER: US 08/598,591  
FILING DATE: 1996-02-12

## ATTORNEY/AGENT INFORMATION:

NAME: Michael S. Tuscan  
REGISTRATION NUMBER: 43,210  
REFERENCE/DOCKET NUMBER: 44921-5053-01-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-739-3000  
TELEFAX: 202-739-3001  
INFORMATION FOR SEQ ID NO: 32:

SEQUENCE CHARACTERISTICS:  
LENGTH: 23 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:



STRAIN: 11ER primer  
SEQUENCE DESCRIPTION: SEQ ID NO: 32:  
US-09-982-828-32

Query Match 11.4% Score 23; DB 9; Length 23;  
Best Local Similarity 100.0%; Pred. No. 0.011;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 154 TGTCAATACCAGTGAAGTAA 176  
|||||  
Db 23 TGTCAATACCAGTGAAGTAA 1

Search completed: June 27, 2003, 06:05:20  
Job time : 159 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 27, 2003, 10:03:36 ; Search time 90 Seconds  
(without alignments)  
2758.288 Million cell updates/sec

Title: US-09-734-672-4  
Perfect score: 9649  
Sequence: 1 MDLSALRVEVQNVINAMQK.....LYCCQLDYLRIPQIPSHY 1863

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

A\_Geneseq\_101002:\*

- 1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*
- 2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*
- 3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*
- 4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*
- 5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*
- 6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*
- 7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*
- 8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*
- 9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*
- 10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*
- 11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*
- 12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*
- 13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*
- 14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*
- 15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*
- 16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*
- 17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*
- 18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*
- 19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*
- 20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*
- 21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*
- 22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*
- 23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	9649	100.0	1863	17	AAR81535	BRCA1 mutant from
2	9649	100.0	1863	19	AAH76099	Human BRCA1 omi2 p
3	9649	100.0	1863	21	AAH24219	Human BRCA1 (omi13)
4	9642	99.9	1863	17	AAR97128	BRCA1, breast and
5	9642	99.9	1863	17	AAR81481	BRCA1, breast and
6	9642	99.9	1863	17	AAR91208	BRCA1, breast and
7	9642	99.9	1863	20	AAV32033	Human BRCA1 protei
8	9642	99.9	2353	22	ABG01696	Novel human diagno
9	9639	99.9	1863	17	AAR81490	BRCA1 mutant from
10	9639	99.9	1863	17	AAR81500	BRCA1 mutant from

11	9639	99.9	1863	17	AAR81522	BRCA1 mutant from
12	9639	99.9	1863	17	AAR81536	BRCA1 mutant from
13	9639	99.9	1863	17	AAR81540	BRCA1 mutant from
14	9638	99.9	1863	17	AAR81529	BRCA1 mutant from
15	9638	99.9	1863	17	AAR81532	BRCA1 mutant from
16	9638	99.9	1863	17	AAR81533	BRCA1 mutant from
17	9638	99.9	1863	17	AAR81534	BRCA1 mutant from
18	9638	99.9	1863	17	AAR81514	BRCA1 mutant from
19	9638	99.9	1863	17	AAR81543	BRCA1 mutant from
20	9638	99.9	1863	17	AAR81545	BRCA1 mutant from
21	9637	99.9	1863	17	AAR81485	BRCA1 mutant from
22	9637	99.9	1863	17	AAR81493	BRCA1 mutant from
23	9637	99.9	1863	17	AAR81505	BRCA1 mutant from
24	9637	99.9	1863	17	AAR81511	BRCA1 mutant from
25	9637	99.9	1863	17	AAR81538	BRCA1 mutant from
26	9636	99.9	1863	17	AAR81486	BRCA1 mutant from
27	9636	99.9	1863	17	AAR81509	BRCA1 mutant from
28	9636	99.9	1863	17	AAR81537	BRCA1 mutant from
29	9636	99.9	1863	17	AAR81542	BRCA1 mutant from
30	9636	99.9	1863	17	AAR81544	BRCA1 mutant from
31	9635	99.9	1863	17	AAR81524	BRCA1 mutant from
32	9635	99.9	1863	17	AAR81546	BRCA1 mutant from
33	9635	99.9	1863	18	AAW26522	Human BRCA1 consen
34	9635	99.9	1863	19	AAW79665	BRCA1 (omi1) prote
35	9635	99.9	1863	19	AAW76100	Human BRCA1 omi13 p
36	9635	99.9	1863	19	AAW76098	Human BRCA1 omi1 p
37	9635	99.9	1863	21	AAH24217	Human BRCA1 (omi1)
38	9635	99.9	1863	21	AAH24218	Human BRCA1 (omi12)
39	9634	99.8	1863	17	AAR81488	BRCA1 mutant from
40	9634	99.8	1863	17	AAR81458	BRCA1 mutant from
41	9634	99.8	1863	17	AAR81539	BRCA1 mutant from
42	9630	99.8	1863	17	AAH99440	BRCA1 allele #8403
43	9630	99.8	1863	17	AAR81497	BRCA1 mutant from
44	9630	99.8	1863	18	AAH10003	Protein encoded by
45	9630	99.8	1863	19	AAW79889	Tumorigenic BRCA1

#### ALIGNMENTS

RESULT 1	
AAH81535	AAH81535 standard; Protein: 1863 AA.
ID	
XX	
AC	AAH81535;
XX	
DT	02-OCT-1996 (first entry)
XX	
DE	BRCA1 mutant from PM06.
XX	
KW	Cancer therapy; breast and ovarian cancer predisposing gene; immunogen;
KW	antibody production; germline alteration; probe; lesion neoplasia; human;
KW	gene therapy; protein replacement therapy; protein mimetic; BRCA1.
XX	
OS	Homo sapiens.
XX	
FT	Key
FT	Misc-difference 871
FT	/note= "p871L"
XX	
PN	MO9605306-A2.
XX	
PD	22-FEB-1996.
XX	
PF	11-AUG-1995;
XX	
XX	95US-0483553.
PR	07-JUN-1995;
PR	12-AUG-1994;
PR	02-SEP-1994;
PR	16-SEP-1994;
PR	29-NOV-1994;
PR	24-MAR-1995;
PR	07-JUN-1995;
PR	95US-0480784.

XX (MYRI-) MYRIAD GENETICS INC.  
 PA (CANC-) CANCER INST.  
 PA (RECH-) CENT RECH DU CHUL.  
 XX  
 PI Durocher F, Emi M, Nakamura Y, Simard J, Shattuck-Eidens DM;  
 DR MPI: 1996-139702/14.  
 DR N-PSDB; AAT17493.  
 XX  
 PT New nucleic acid and polypeptide for mutant or polymorphic BRCA1  
 PT gene - for diagnosis and therapy of human breast and ovarian cancer  
 PT and for diagnosing pre-disposition to these cancers  
 PS  
 PS Claim 1: 218bp; English.  
 XX  
 CC AAR81483-R81497 and AAR81499-R81546 represent mutations of the protein  
 CC encoded by the human breast and ovarian cancer predisposing gene (BRCA1)  
 CC (see AAR81481 for wild type protein). These mutations can be used as  
 CC immunogens for antibody production. The mutant BRCA1 genes encoding  
 CC these sequences have at least 1 mutation or polymorphism in comparison  
 CC to the wild type cDNA (see AAT17438 for wild type). By detecting a  
 CC germline alteration in the wild type BRCA1 gene, a predisposition for  
 CC breast and ovarian cancer can be diagnosed. In one method, BRCA1 mRNA  
 CC isolated from a tissue sample from a subject has a probe, corresponding  
 CC to a fragment of the cDNA encoding the wild type BRCA1 sequence (or an  
 CC allele-specific probe for a mutation of it), added to it. The conditions  
 CC allow for hybridization of the probe to the mRNA, and any hybridisation  
 CC which occurs is detected. Alternatively the BRCA1 gene in the tissue  
 CC sample is isolated, and a shift in electrophoretic mobility of single  
 CC stranded DNA from the sample on a non-denaturing polyacrylamide gel  
 CC indicates a mutation. These methods of detection can also diagnose a  
 CC lesion neoplasia associated with the BRCA1 locus. The methods may be  
 CC used in gene therapy, protein replacement therapy and protein mimetics,  
 CC and may be used to screen for drugs in cancer therapy.  
 XX  
 XX Sequence 1863 AA:  
 SQ  
 Query Match 100.0%; Score 9649; DB 17; Length 1863;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1863; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 481 LIIGAFVTEPOIIQERPLTNKLRKRRTPTSGLPEDFIKADLAVOKTPREMIINOCTNOTE 540  
 DB 481 LIIGAFVTEPOIIQERPLTNKLRKRRTPTSGLPEDFIKADLAVOKTPREMIINOCTNOTE 540  
 OY 541 QNGOVNMITNSGHNKTGDSIONEKNPNDEISLEKESAFTKAEDISSISIMLELNI 600  
 DB 541 QNGOVNMITNSGHNKTGDSIONEKNPNDEISLEKESAFTKAEDISSISIMLELNI 600  
 OY 601 HNSKAPKKNLRRKSSRTHIALELVYRNLSPNCTEQIDSCSSSEIKKKKNOMV 660  
 DB 601 HNSKAPKKNLRRKSSRTHIALELVYRNLSPNCTEQIDSCSSSEIKKKKNOMV 660  
 OY 601 HNSKAPKKNLRRKSSRTHIALELVYRNLSPNCTEQIDSCSSSEIKKKKNOMV 660  
 DB 601 HNSKAPKKNLRRKSSRTHIALELVYRNLSPNCTEQIDSCSSSEIKKKKNOMV 660  
 OY 661 RHSRNLQIMEGKEPATGAKKSNKPNQTSKRHSDPFPELKLINAGSFSTKSNTSELKE 720  
 DB 661 RHSRNLQIMEGKEPATGAKKSNKPNQTSKRHSDPFPELKLINAGSFSTKSNTSELKE 720  
 OY 721 FVNSLPREBEKEKLETVKYSNNAEDPKDMLSGERVLOTERSVESSISLVPGTDTGYO 780  
 DB 721 FVNSLPREBEKEKLETVKYSNNAEDPKDMLSGERVLOTERSVESSISLVPGTDTGYO 780  
 OY 781 ETSISLEVSTLGRKAKTFPNKCVSQCAFENPKCLIGCSKDNNDTEGFRKPLGHEVNS 840  
 DB 781 ETSISLEVSTLGRKAKTFPNKCVSQCAFENPKCLIGCSKDNNDTEGFRKPLGHEVNS 840  
 OY 841 RETSIEMESELDAOYLONTFKYSKROSFALFNPNGNAEBECATFSAHSGSLKKOSPKYT 900  
 DB 841 RETSIEMESELDAOYLONTFKYSKROSFALFNPNGNAEBECATFSAHSGSLKKOSPKYT 900  
 OY 901 FECEOKENEGKNESNIKRPQVTNITAGFVYVQOKDPVDMAKCSIKGSRCLSSQFNG 960  
 DB 901 FECEOKENEGKNESNIKRPQVTNITAGFVYVQOKDPVDMAKCSIKGSRCLSSQFNG 960  
 OY 961 NENGLITPNKHGLQNLNRYRPLPFIKSPYTKCKNLEENPEEHSNPREEMKENENTP 1020  
 DB 961 NENGLITPNKHGLQNLNRYRPLPFIKSPYTKCKNLEENPEEHSNPREEMKENENTP 1020  
 OY 1021 STVSTISRNNIRENVEKEASSNINEVSGSTNEVGSSINEIGSDENIOAELGRNRGPTL 1080  
 DB 1021 STVSTISRNNIRENVEKEASSNINEVSGSTNEVGSSINEIGSDENIOAELGRNRGPTL 1080  
 OY 1081 NAMRLGLVLOPEVYKOSLPQSNCKKPRFKOEYEBVQVNTDFSPYLLISDLNEDPMGSS 1140  
 DB 1081 NAMRLGLVLOPEVYKOSLPQSNCKKPRFKOEYEBVQVNTDFSPYLLISDLNEDPMGSS 1140  
 OY 1141 HASOVCEPDPDDLDDGEIKEDTSPFANDIKESAVFSKSVQGELSRSPFTHTHLAQ 1200  
 DB 1141 HASOVCEPDPDDLDDGEIKEDTSPFANDIKESAVFSKSVQGELSRSPFTHTHLAQ 1200  
 OY 1201 GYRGAKKLESSEENLSEDEDELPCFQHLIFGVNNIPQOSTRHSTVATFELSKNTENTL 1260  
 DB 1201 GYRGAKKLESSEENLSEDEDELPCFQHLIFGVNNIPQOSTRHSTVATFELSKNTENTL 1260  
 OY 1261 LSLKNSLNDCSNOVILAKASQEHHLSEETKCSASLFSQCSSELEDITANTNTQDPFLIGS 1320  
 DB 1261 LSLKNSLNDCSNOVILAKASQEHHLSEETKCSASLFSQCSSELEDITANTNTQDPFLIGS 1320  
 OY 1321 SKOMRHQSESQVGLSDKELVSDDERGRTGLEENNOBESMNSNLGEAASGCESESTVSE 1380  
 DB 1321 SKOMRHQSESQVGLSDKELVSDDERGRTGLEENNOBESMNSNLGEAASGCESESTVSE 1380  
 OY 1381 DCSGLSSQSDILFTQORDTPQOHLIKLOOEAMELEAVLVOHQSOPSNSYPSITISSSALE 1440  
 DB 1381 DCSGLSSQSDILFTQORDTPQOHLIKLOOEAMELEAVLVOHQSOPSNSYPSITISSSALE 1440  
 OY 1441 DLBNPEOSTSEKAVLTSQKSSEYPIQONPEGLSADKFEVYSAADSSTSKNKEPGEVBSRPSK 1500  
 DB 1441 DLBNPEOSTSEKAVLTSQKSSEYPIQONPEGLSADKFEVYSAADSSTSKNKEPGEVBSRPSK 1500  
 OY 1501 CPSLDDBRWYMHSCSGSLQNRNRPQSEELIKYVDVEEQOLEBEGPHDLTFTSTLPLQODLEG 1560  
 DB 1501 CPSLDDBRWYMHSCSGSLQNRNRPQSEELIKYVDVEEQOLEBEGPHDLTFTSTLPLQODLEG 1560

QY 1561 TPYLESGISLSPDDESPEDSRAPESARVGNIPSSSTALKVQDLKVAESASQSPAHAHTT 1620  
 DB 1561 TPYLESGISLSPDDESPEDSRAPESARVGNIPSSSTALKVQDLKVAESASQSPAHAHTT 1620  
 QY 1621 DTAGYANMESVSREKRELTASTERVKRMSVYSGITPEEFMLVYKFAKHHITLTNLI 1680  
 DB 1621 DTAGYANMESVSREKRELTASTERVKRMSVYSGITPEEFMLVYKFAKHHITLTNLI 1680  
 QY 1681 TEETHVHMKTDAEFCEERTLKFELGIAGKMWYVFWYQSIKERKMLNEHPEVARGDY 1740  
 DB 1681 TEETHVHMKTDAEFCEERTLKFELGIAGKMWYVFWYQSIKERKMLNEHPEVARGDY 1740  
 QY 1741 VNGRNHOGPKRARESDOKIFRGLIEICCYGPTNMPDQLEMMVQLOGASVARELSFTL 1800  
 DB 1741 VNGRNHOGPKRARESDOKIFRGLIEICCYGPTNMPDQLEMMVQLOGASVARELSFTL 1800  
 QY 1801 GTGVHPIVYVQAPAMWEDNCFHAIIGOMCEAPVYTRREVLDVALYQCELDITLIPQIPH 1860  
 DB 1801 GTGVHPIVYVQAPAMWEDNCFHAIIGOMCEAPVYTRREVLDVALYQCELDITLIPQIPH 1860  
 QY 1861 SHY 1863  
 DB 1861 SHY 1863

RESULT 2  
 AAM76099  
 ID AAM76099 standard; Protein; 1863 AA.  
 XX

AC AAM76099;  
 DT 18-NOV-1998 (first entry)  
 DE Human BRCA1 cml2 protein.

KW BRCA1: cml2; human; breast and ovarian cancer predisposing gene;  
 KW polymorphism; susceptibility; anti-oncogene; tumour suppressor;  
 KW chromosome 17q.

XX Homo sapiens.

PN US5750400-A.

PD 12-MAY-1998.

PF 12-FEB-1997; 9705-0798691.

PR 12-FEB-1996; 96US-0598591.

PR 12-FEB-1997; 9705-0798691.

XX (ONCO-) ONCORMED INC.

PI Allen AC, Alvares CP, Critz BS, Murphy PD, Olson SJ;  
 PI Schelter DB, Zeng B;

DR WPI: 1998-296774/26.

DR N-PSDB; AAV46449.

PT BRCA1 cml gene coding sequences - useful for distinguishing between  
 PT polymorphisms and mutation(s) in the screening for disposition to  
 PT breast or ovarian cancer

XX disclosure; Column 47-56; 54pp; English.

XX This sequence represents the human BRCA1 (breast and ovarian cancer  
 XX predisposing gene) cml2 gene. This sequence and polymorphic variations of  
 CC this sequence are useful for the identification of an individual who may  
 CC or may not have an increased susceptibility to breast or ovarian cancer.  
 CC The sequences used identify gene changes which are due to polymorphisms,  
 CC rather than cancer causing mutations. BRCA1 is an anti-oncogene (tumour  
 CC suppressor) which is involved in genetic inheritance of cancers,  
 CC especially breast and ovarian cancer. It is found at human chromosome 17q  
 CC which is known to be linked to cancer susceptibility, especially breast

CC cancer. Cells containing a mutation in this gene lose the wild-type  
 CC function of BRCA1 and are more susceptible to cancers.

XX Sequence 1863 AA:

Query Match 100.0%; Score 9649; DB 19; Length 1863;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1863; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDLSALRVEEVQVYVNMOKILBEPICLLEKEVSTKCHIRKFCFMLKLNOKKPSQ 60  
 DB 1 MDLSALRVEEVQVYVNMOKILBEPICLLEKEVSTKCHIRKFCFMLKLNOKKPSQ 60  
 QY 61 CPLKNDITKRSLOESTRFQOLVBEELIKICAFOLDLGEVANSYNPAKKENSPBLKD 120  
 DB 61 CPLKNDITKRSLOESTRFQOLVBEELIKICAFOLDLGEVANSYNPAKKENSPBLKD 120  
 QY 121 EVSIIOSMGYRNRAKRLQSEPNPSLOETSLVSOLNLTGVRTLRTKORIOPQTSVI 180  
 DB 121 EVSIIOSMGYRNRAKRLQSEPNPSLOETSLVSOLNLTGVRTLRTKORIOPQTSVI 180  
 QY 181 ELGSDSEDTVNAKATYCSVGDDELQTTPOGTREISLDSAKKACFSEDTVNTNHHQ 240  
 DB 181 ELGSDSEDTVNAKATYCSVGDDELQTTPOGTREISLDSAKKACFSEDTVNTNHHQ 240  
 QY 241 PSNNDLNTTEKRAAERHPEKYOGSSVNLHVEPCGTHTASSLOHENSLLTKDRMNV 300  
 DB 241 PSNNDLNTTEKRAAERHPEKYOGSSVNLHVEPCGTHTASSLOHENSLLTKDRMNV 300  
 QY 301 KAEFCNKSQPGIARSOHNWAGSKETCNDRTPTSTKRYDLNADPLCEKREWNKOKLPC 360  
 DB 301 KAEFCNKSQPGIARSOHNWAGSKETCNDRTPTSTKRYDLNADPLCEKREWNKOKLPC 360  
 QY 361 SENPROTEDVPTLNTSSICKVNMFSRSDLLGSDSHGSESNKAVADVILNEVD 420  
 DB 361 SENPROTEDVPTLNTSSICKVNMFSRSDLLGSDSHGSESNKAVADVILNEVD 420  
 QY 421 EYSGSSSEKIDLASDPHEALICKSERVHSKVSSENIEDKIFGTYRKKAALPNLSHTEN 480  
 DB 421 EYSGSSSEKIDLASDPHEALICKSERVHSKVSSENIEDKIFGTYRKKAALPNLSHTEN 480  
 QY 481 LITGAFTPEOITIOERPLTNKLRKRRTSGLHPEDEIKRADLAVOKTPEMINOQNOTE 540  
 DB 481 LITGAFTPEOITIOERPLTNKLRKRRTSGLHPEDEIKRADLAVOKTPEMINOQNOTE 540  
 QY 541 QNGOVNNTNSGHENKTGDSIONEKPNPIESLEKESAKTAAEPLSSISNMELENT 600  
 DB 541 QNGOVNNTNSGHENKTGDSIONEKPNPIESLEKESAKTAAEPLSSISNMELENT 600  
 QY 601 HNSKAPKKNRLRRKSSRTHTALELVYSRNLSPPNCTELOIDSCSSSEELKKKKYNOMP 660  
 DB 601 HNSKAPKKNRLRRKSSRTHTALELVYSRNLSPPNCTELOIDSCSSSEELKKKKYNOMP 660  
 QY 661 RHSRNLQLMGKREPATGAKKSNKPNEDQSKRHSDTFPELKLTNAPGSFTKCSNTSELE 720  
 DB 661 RHSRNLQLMGKREPATGAKKSNKPNEDQSKRHSDTFPELKLTNAPGSFTKCSNTSELE 720  
 QY 721 FVNPSLPREKEKLELVYKVSNNNAEDPKDMLSGERVLQFERVSESSISLVGCTDYGTO 780  
 DB 721 FVNPSLPREKEKLELVYKVSNNNAEDPKDMLSGERVLQFERVSESSISLVGCTDYGTO 780  
 QY 781 ESISLLEVTGLGAKTEPNKCVSOCAAFENPKGLIHGCSKDNNDTGFYTPGLGHEVNH 840  
 DB 781 ESISLLEVTGLGAKTEPNKCVSOCAAFENPKGLIHGCSKDNNDTGFYTPGLGHEVNH 840  
 QY 841 RETSIEMESELDAOYLQNTFKVSKROSFALFSPNPNAAEECATFSAHSGSLKQSPKVT 900  
 DB 841 RETSIEMESELDAOYLQNTFKVSKROSFALFSPNPNAAEECATFSAHSGSLKQSPKVT 900  
 QY 901 FECEOKREENGKNKESNIPVQYVNTAGPVPVQOKKPVDNACSTIKGSRFLSSQFRG 960  
 DB 901 FECEOKREENGKNKESNIPVQYVNTAGPVPVQOKKPVDNACSTIKGSRFLSSQFRG 960

```

OY 961 NETGLITPNKHGLLONPYRIPLPLPIKSFVKTKCKKNLLEENFEHSMSPEREMGNENTP 1020
DB 961 NETGLITPNKHGLLONPYRIPLPLPIKSFVKTKCKKNLLEENFEHSMSPEREMGNENTP 1020
OY 1021 STVSTISRNINIRENVKKEASSSNINEVSGSTNEVGSINIEIGSSDENIQAELOCRNGPKL 1080
DB 1021 STVSTISRNINIRENVKKEASSSNINEVSGSTNEVGSINIEIGSSDENIQAELOCRNGPKL 1080
OY 1081 NAMRLGVQPEVYKQSLPGSNCKHPEIKKOEYEEVQVNTDFESYLLSDNLEOPMGSS 1140
DB 1081 NAMRLGVQPEVYKQSLPGSNCKHPEIKKOEYEEVQVNTDFESYLLSDNLEOPMGSS 1140
OY 1141 HASQVCESTPDDLDDGEIKEDTSFAENDIKESSAAVFSKSVOKGELSRSPPTHTHLA 1200
DB 1141 HASQVCESTPDDLDDGEIKEDTSFAENDIKESSAAVFSKSVOKGELSRSPPTHTHLA 1200
OY 1201 GYRGAKKLSESEENLSEDEELPCFQHLFGKVNIPISOSTRSTVATECLSKNTEENL 1260
DB 1201 GYRGAKKLSESEENLSEDEELPCFQHLFGKVNIPISOSTRSTVATECLSKNTEENL 1260
OY 1261 LSLKNSLNDCSNOVILAKASQEHHLSEETKCSASLFSQCELEDELITANTNTODPFLIGS 1320
DB 1261 LSLKNSLNDCSNOVILAKASQEHHLSEETKCSASLFSQCELEDELITANTNTODPFLIGS 1320
OY 1321 SKQMRHOSQSGVGLSDKELVSDDERGTCGLEENNOCESGDSNLCGAASGCESETSVSE 1380
DB 1321 SKQMRHOSQSGVGLSDKELVSDDERGTCGLEENNOCESGDSNLCGAASGCESETSVSE 1380
OY 1381 DCSGSSOSDILITTOORDTMOHNLIKLOQEMAELEVLROHGSQPSNSTPSTISDSALE 1440
DB 1381 DCSGSSOSDILITTOORDTMOHNLIKLOQEMAELEVLROHGSQPSNSTPSTISDSALE 1440
OY 1441 DLRNPEQSTSEKAVILTSOKSEYPISQNEBGLSADKFEVYSAADSTSKNKEPGEVRSRSPK 1500
DB 1441 DLRNPEQSTSEKAVILTSOKSEYPISQNEBGLSADKFEVYSAADSTSKNKEPGEVRSRSPK 1500
OY 1501 CPSLDDRWYMHSCSGSLQNRNRPQOEELIKYVDVEEQULEESGPHDLTETSTYLRDDELG 1560
DB 1501 CPSLDDRWYMHSCSGSLQNRNRPQOEELIKYVDVEEQULEESGPHDLTETSTYLRDDELG 1560
OY 1561 TPYLESGISLFSDDPESDSEDRAPASAVGNIPSTSTALKVQOLVAESAQSPAAAHHT 1620
DB 1561 TPYLESGISLFSDDPESDSEDRAPASAVGNIPSTSTALKVQOLVAESAQSPAAAHHT 1620
OY 1621 DTAGYNAMEESVSRKPELTASTERYNKRMSVVSGLTPEEFMLVYKFAKHHITLTNLI 1680
DB 1621 DTAGYNAMEESVSRKPELTASTERYNKRMSVVSGLTPEEFMLVYKFAKHHITLTNLI 1680
OY 1681 TEETHHYVVKTPAEVCEKTLKYPFLGIAGCKWVVSFWVTQSTKEKMLNEHDFEVRGV 1740
DB 1681 TEETHHYVVKTPAEVCEKTLKYPFLGIAGCKWVVSFWVTQSTKEKMLNEHDFEVRGV 1740
OY 1741 VNGRHHQGPKRARESQDRKIFRGLEICCYGPTNMPDTDOLEMMVQCGASVKELESSTFL 1800
DB 1741 VNGRHHQGPKRARESQDRKIFRGLEICCYGPTNMPDTDOLEMMVQCGASVKELESSTFL 1800
OY 1801 GTGVHPPIVVOGDANTEDNGFHAICOMCEAPVVTREWLDSVALYQOQELDTYLLPQIPH 1860
DB 1801 GTGVHPPIVVOGDANTEDNGFHAICOMCEAPVVTREWLDSVALYQOQELDTYLLPQIPH 1860
OY 1861 SHY 1863
DB 1861 SHY 1863

```

RESULT 3  
 AAB24219  
 ID AAB24219 standard; Protein: 1863 AA.  
 XX AAB24219;  
 XX 07-FEB-2001 (first entry)  
 XX

```

DE Human BRCA1 (om13) protein sequence SEQ ID NO:6.
XX
XX Human: BRCA1: chromosome 17: 17q21; breast cancer; ovarian cancer;
KW gene therapy; diagnosis; cytostatic; genetic susceptibility; mutation;
KW polymorphism; identification.
OS Homo sapiens.
XX
XX US6130322-A.
XX
XX 10-OCT-2000.
XX
XX 06-MAY-1998; 9805-0074476.
XX
XX 12-FEB-1996; 9605-0598591.
XX 12-DEC-1997; 9705-0798691.
XX
XX (GENE-) GENE LOGIC INC.
XX
XX Zeng B, Thiruber D, Olson SJ, Alvares CP, Allen ACP, Murphy PD;
PI Critz BS;
XX
XX WPI: 2000-646756/62.
XX
XX N-PSDB: AAC60795.
XX
XX Example 4; Column 65-74; 56pp; English.
XX
XX AAC60793 to AAC60795 encode the human BRCA1 (om1-3) proteins given in
CC AAB24217 to AAB24219 respectively. BRCA1 is found on chromosome 17
CC mapping to position 17q21. The BRCA1 (om12) coding sequence is
CC specifically claimed in the present invention. The BRCA1 (om12) coding
CC sequence is useful in gene therapy, especially for preventing or treating
CC breast or ovarian cancer. It is also useful for diagnosing or monitoring
CC breast or ovarian cancer. Furthermore, the BRCA1 (om12) coding sequence
CC is useful for: (a) identifying individuals having BRCA1 gene mutations
CC and having an increased genetic susceptibility to breast or ovarian
CC cancer, or identifying a mutation that increases the genetic
CC susceptibility to breast or ovarian cancer; (b) avoiding
CC misinterpretation of polymorphisms found in the BRCA1 gene; (c)
CC determining the presence of a previously unknown mutation in the BRCA1
CC gene; (d) probing a human sample of the BRCA1 gene by allele to determine
CC the presence of either polymorphic alleles or mutations; and (e)
CC performing diagnosis with a reagent derived from the BRCA1 (om1) cDNA
CC sequence. AAC60796 to AAC60861 represent PCR primers for the BRCA1 gene,
CC which are used in an example from the present invention.
XX
XX
SQ Sequence 1863 AA:
Query Match 100.0%; Score 9649; DB 21; Length 1863;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1863; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 MDLSALRVEEYQNVINAMQKILECPICELKEPSTKCDHIFCFMILKLNQKGPQ 60
DB 1 MDLSALRVEEYQNVINAMQKILECPICELKEPSTKCDHIFCFMILKLNQKGPQ 60
OY 61 CPLCKNDITKRSLOESTRESQLVEELIKITCAFQDLDTGLEVANSYNFAKKENNSPEHLKD 120
DB 61 CPLCKNDITKRSLOESTRESQLVEELIKITCAFQDLDTGLEVANSYNFAKKENNSPEHLKD 120
OY 121 EVSIIQSKGYRNKARKRLQSEPNLSQETSLSVQSLNGLGIVRTLRTQRIQPKTSYI 180
DB 121 EVSIIQSKGYRNKARKRLQSEPNLSQETSLSVQSLNGLGIVRTLRTQRIQPKTSYI 180
OY 181 ELGSDSSEDTVKKATYCSVGOEQLQITPOGRDSTLSDSKAKACESEFDVNTTEHQ 240
DB 181 ELGSDSSEDTVKKATYCSVGOEQLQITPOGRDSTLSDSKAKACESEFDVNTTEHQ 240

```

```

QY 241 PSNDLTTTEKRAAEHREPEKYQSSVSNLHVEPCGTNTHASSLOHENSLLTLDKDMNVE 300
D 241 PSNDLTTTEKRAAEHREPEKYQSSVSNLHVEPCGTNTHASSLOHENSLLTLDKDMNVE 300
QY 301 KAECNKSQOPGLARSOHNRWAGSKETCNDRPTSTBKVDLNDPLCERENMKOKLPC 360
D 301 KAECNKSQOPGLARSOHNRWAGSKETCNDRPTSTBKVDLNDPLCERENMKOKLPC 360
QY 361 SENPRDEDPWMLTINSIOKVNMFPISRDELLSDSDSHOSESNAKVADVLDVLEVD 420
D 361 SENPRDEDPWMLTINSIOKVNMFPISRDELLSDSDSHOSESNAKVADVLDVLEVD 420
QY 421 EYSSGSEKIDLLASDPHEALICKSERVHKSVESNIEDKIFGKTYRRKASLPLNSHVTEN 480
D 421 EYSSGSEKIDLLASDPHEALICKSERVHKSVESNIEDKIFGKTYRRKASLPLNSHVTEN 480
QY 481 LIIGAFTPEPOITIOERPLTNKIKRKRPTSGLHPEDETIKKADLAVQTPREKINGNOTE 540
D 481 LIIGAFTPEPOITIOERPLTNKIKRKRPTSGLHPEDETIKKADLAVQTPREKINGNOTE 540
QY 541 QNGOVANITNSGHEKNTKQDSIQNEKNPNPIESLEKESAFTKAPETISSISNNLELNI 600
D 541 QNGOVANITNSGHEKNTKQDSIQNEKNPNPIESLEKESAFTKAPETISSISNNLELNI 600
QY 601 HNSKAPKKNRLRKSSTRHIALELVVSRNLSPNCTELQIDSCSSSEETIKKKKYNOMPV 660
D 601 HNSKAPKKNRLRKSSTRHIALELVVSRNLSPNCTELQIDSCSSSEETIKKKKYNOMPV 660
QY 661 RHSNLIOLMEGKEPATGAKKSNNKNEOTSKRHSDTPELKLTAAPBSFTKCSMTSELKE 720
D 661 RHSNLIOLMEGKEPATGAKKSNNKNEOTSKRHSDTPELKLTAAPBSFTKCSMTSELKE 720
QY 721 FVNPSPREKEEKELETVKVSNNNAEDPKDMLSGSERVLOTREVSSESSISLVPSTDVGTQ 780
D 721 FVNPSPREKEEKELETVKVSNNNAEDPKDMLSGSERVLOTREVSSESSISLVPSTDVGTQ 780
QY 781 ESISLEFVSTLGAKTEPNKCVSOCAAFENPKGLIHGCSKDNDRNDTEGFKYPLGHEVNH 840
D 781 ESISLEFVSTLGAKTEPNKCVSOCAAFENPKGLIHGCSKDNDRNDTEGFKYPLGHEVNH 840
QY 841 RETSIEMESELDAQOYONTFKVSKROSFAFNSRNGAEECAFTFSHSGSLKQSPKVT 900
D 841 RETSIEMESELDAQOYONTFKVSKROSFAFNSRNGAEECAFTFSHSGSLKQSPKVT 900
QY 901 FECGOKRENGKNSNIKPQVYNITAGFPVYGOKDPVDNAKCSIGSGRFLCSSOPRG 960
D 901 FECGOKRENGKNSNIKPQVYNITAGFPVYGOKDPVDNAKCSIGSGRFLCSSOPRG 960
QY 961 NETGLITPNKHGLLQNPYRIPLPEPIKSFVYTKCKKNLLENFEBHSMSPERENGNEIP 1020
D 961 NETGLITPNKHGLLQNPYRIPLPEPIKSFVYTKCKKNLLENFEBHSMSPERENGNEIP 1020
QY 1021 STVSTIRNNIREVNEFEASSNINEVGSSTNEVGSSINELGSSDENIQAELGKRNRPKL 1080
D 1021 STVSTIRNNIREVNEFEASSNINEVGSSTNEVGSSINELGSSDENIQAELGKRNRPKL 1080
QY 1081 NAMRLGLVLOPEYVKOSLPSNCKHPETIKOEVEVQVNTDPSPLIDNLEQPMGSS 1140
D 1081 NAMRLGLVLOPEYVKOSLPSNCKHPETIKOEVEVQVNTDPSPLIDNLEQPMGSS 1140
QY 1141 HASOVCEPDDLLDDEIKEDTSFAENDIKESSAVFSKSVOKGELSRSPPTHTHIAQ 1200
D 1141 HASOVCEPDDLLDDEIKEDTSFAENDIKESSAVFSKSVOKGELSRSPPTHTHIAQ 1200
QY 1201 GYRRGAKKLSSSEENLSSDEBELPCFQHLLEFGKYNNTIPSSOSTRSTYATATECLSKNTEENL 1260
D 1201 GYRRGAKKLSSSEENLSSDEBELPCFQHLLEFGKYNNTIPSSOSTRSTYATATECLSKNTEENL 1260
QY 1261 LSLKNSLNDSCNOYITLAKASQEHHLSEFTKCSASLFFSSQSELEDLTANNTQDPLIGS 1320
D 1261 LSLKNSLNDSCNOYITLAKASQEHHLSEFTKCSASLFFSSQSELEDLTANNTQDPLIGS 1320
QY 1321 SKOMRHOSQGVGLSDKELVSDDEERGTLLENNOEQSMDSNLGRASGCESETSVSE 1380

```

```

D 1321 SKOMRHOSQGVGLSDKELVSDDEERGTLLENNOEQSMDSNLGRASGCESETSVSE 1380
QY 1381 DCSGLSSOSDILITQOQDPTMOHNLIKLOEAAELEAVLEOHGSPSNSYPIIISDSSALE 1440
D 1381 DCSGLSSOSDILITQOQDPTMOHNLIKLOEAAELEAVLEOHGSPSNSYPIIISDSSALE 1440
QY 1441 DLARPQESTSEKAVLTQKSSEEPYISQNPBGLSADKEFVSADSSSTSNKKEPVGRSSPSK 1500
D 1441 DLARPQESTSEKAVLTQKSSEEPYISQNPBGLSADKEFVSADSSSTSNKKEPVGRSSPSK 1500
QY 1501 CPSLDDRWYMHSCGSLQNRNYPQOEBELIKVYDVEEQOLESGPHDLTETSYLRDQLEG 1560
D 1501 CPSLDDRWYMHSCGSLQNRNYPQOEBELIKVYDVEEQOLESGPHDLTETSYLRDQLEG 1560
QY 1561 TPYLESGISLPSDDPESDPSDRAPESARGNIPSSSTALKVPOLKVAESQSPAANTT 1620
D 1561 TPYLESGISLPSDDPESDPSDRAPESARGNIPSSSTALKVPOLKVAESQSPAANTT 1620
QY 1621 DTAGYNAMESVSREKEBELTASTERVNKRMSMVVSGLTPEBFMLVYFARKHITLNL 1680
D 1621 DTAGYNAMESVSREKEBELTASTERVNKRMSMVVSGLTPEBFMLVYFARKHITLNL 1680
QY 1681 TEETHVVMKTDAEFVCEERTLYFLGIAGCKWVSVYFWVTQSIKERKMLNEHDEVRGCV 1740
D 1681 TEETHVVMKTDAEFVCEERTLYFLGIAGCKWVSVYFWVTQSIKERKMLNEHDEVRGCV 1740
QY 1741 VNGRNHOGPKRARSORRTFRGLEICCYGFTWMPPTDOLBMWVOLGASVYKLSFTL 1800
D 1741 VNGRNHOGPKRARSORRTFRGLEICCYGFTWMPPTDOLBMWVOLGASVYKLSFTL 1800
QY 1801 GTGVHPIVVQPDAMTEWDNGFHAIGOMCEAPVYVRVWVLDVSVALYOCQELDTYLP 1860
D 1801 GTGVHPIVVQPDAMTEWDNGFHAIGOMCEAPVYVRVWVLDVSVALYOCQELDTYLP 1860
QY 1861 SHY 1863
D 1861 SHY 1863

RESULT 4
AAR97128
ID AAR97128 standard; Protein: 1863 AA.
XX AC AAR97128:
XX DT 19-NOV-1996 (first entry)
XX DE BRCA1, breast and ovarian cancer susceptibility gene product.
XX KW BRCA1; breast cancer; ovary cancer; predisposing gene; diagnosis;
susceptibility gene; prognosis; gene therapy.
XX OS Homo sapiens.
XX PN W09605308-A1.
XX PD 22-FEB-1996.
XX PF 11-AUG-1995; 95WO-0510220.
XX PR 07-JUN-1995; 95US-0488011.
PR 12-AUG-1994; 94US-0289221.
PR 02-SEP-1994; 94US-0300266.
PR 16-SEP-1994; 94US-0308104.
PR 29-NOV-1994; 94US-0348824.
PR 24-MAR-1995; 95US-0409305.
PR 07-JUN-1995; 95US-0483554.
PR 07-JUN-1995; 95US-0487002.
XX PA (MYRI-) MYRIAD GENETICS INC.
PA (UTAH) UNIV UTAH RES FOUND.
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.

```

XX Futreal PA, Goldgar DE, Harshman KD, Kamb A, Miki Y;  
PI Shattuck-Eidens DM, Skolnick MH, Swenson J, Tavtigian SV;  
PI Mismen RM;  
XX WPI, 1996-139704/14.  
DR N-PSDB; AAT32601.  
XX  
PT New method for diagnosing a predisposition to breast and ovarian  
PT cancer - by detecting a germline alteration in the BRCA1 gene or  
PT gene regulatory sequence; for gene therapy and to screen for drugs  
XX  
PS Claim 12: Page 108-117; 200pp: English.

This is the amino acid sequence of the protein encoded by the breast and  
ovarian cancer susceptibility gene, BRCA1. Four kindred families  
provided genetic evidence for localisation of BRCA1 to a sufficiently  
small region for the application of positional cloning strategies. A  
detailed map of transcripts was developed for the region of 17q21 between  
CC D17S1321 and D17S1324. A combination of sequences obtd. from cDNA  
CC clones, hybrid-selected sequences and PCR prods. allowed construction of  
CC a composite full-length BRCA1 cDNA (see AAT32612 for genomic sequence).  
CC The BRCA1 cDNA revealed an open reading frame encoding a protein of  
CC 208 KD. The protein also contains a sequence near the amino terminus  
CC which had homology to zinc finger domains, esp. the sequence contains  
CC Cys and His residues present as a C3HC4 zinc finger consensus motif.  
CC The isolated cDNA is used in methods for either diagnosis of the  
CC predisposition to cancer (partic. breast and ovarian cancer), or for the  
CC diagnosis or prognosis of cancer, and also in gene-based therapies  
CC directed at cancer cells.

XX Sequence 1863 AA:

Query Match 99.9%; Score 9642; DB 17; Length 1863;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1862; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDLSALRVEEYQVNIAMQKILICPILCELIKPEVSTKCDHIFCKFCMLLNQKKGPSQ 60  
DB 1 MDLSALRVEEYQVNIAMQKILICPILCELIKPEVSTKCDHIFCKFCMLLNQKKGPSQ 60  
QY 61 CPLCKNDITKRSLQESTRESQVLEELKTIICAFQDQTGLEAYNSYNPAKKENNSPELKD 120  
DB 61 CPLCKNDITKRSLQESTRESQVLEELKTIICAFQDQTGLEAYNSYNPAKKENNSPELKD 120  
QY 121 EVSIISQMGYRNRAKRLQSEPNPSLQETSLSVQLSNLGTVRTLRKQKIQOKTSVYI 180  
DB 121 EVSIISQMGYRNRAKRLQSEPNPSLQETSLSVQLSNLGTVRTLRKQKIQOKTSVYI 180  
QY 181 ELGSDSSEDTYVKATYCSVGDDELLQITPGTRDEISLDSAKKACFESETDVTNTEHQ 240  
DB 181 ELGSDSSEDTYVKATYCSVGDDELLQITPGTRDEISLDSAKKACFESETDVTNTEHQ 240  
QY 241 PENNDLNTTEKRAEHPKRYGSSVSNLHVEPCGNTNTHASSLQHEHNSLLTKDRNVE 300  
DB 241 PENNDLNTTEKRAEHPKRYGSSVSNLHVEPCGNTNTHASSLQHEHNSLLTKDRNVE 300  
QY 301 KAEPCKSKQOPGLARQOHNRNMAKSEKTCNDRTPTSTKRYVDLNAADPLCEKKNKOKLPC 360  
DB 301 KAEPCKSKQOPGLARQOHNRNMAKSEKTCNDRTPTSTKRYVDLNAADPLCEKKNKOKLPC 360  
QY 361 SENPRDTEVPWITLNSIQKVNEMFSRDELLGSDSDSHDGESESNKAVADVLDVNEVD 420  
DB 361 SENPRDTEVPWITLNSIQKVNEMFSRDELLGSDSDSHDGESESNKAVADVLDVNEVD 420  
QY 421 EYSGSSEKTDLLASOPHEALICKSEKRVHSHKSVSNEDKIFGKTYKKASLPMLSHVTEN 480  
DB 421 EYSGSSEKTDLLASOPHEALICKSEKRVHSHKSVSNEDKIFGKTYKKASLPMLSHVTEN 480  
QY 481 LITGAFVTEPQITQIERPLTNKLRKRRTSGLHPDEFIKKADLAVOKTPMIMQGNQTE 540  
DB 481 LITGAFVTEPQITQIERPLTNKLRKRRTSGLHPDEFIKKADLAVOKTPMIMQGNQTE 540

QY 541 QNGQVNNITNSGHENKTKGDSIQNEKNPNPIESLEKESAFKTKAEPSSISSNMELEUNI 600  
DB 541 QNGQVNNITNSGHENKTKGDSIQNEKNPNPIESLEKESAFKTKAEPSSISSNMELEUNI 600  
QY 601 HNSKAPKKNNLRKRSSTRHIALELVYSRNLSPNCTELQIDSCSSEIKKKKNQMPV 660  
DB 601 HNSKAPKKNNLRKRSSTRHIALELVYSRNLSPNCTELQIDSCSSEIKKKKNQMPV 660  
QY 661 RHRNQLQMGKEPAPGAKKSNKPNQOTSKRHSDTFPEELKLTNMGSTKCSNTEKE 720  
DB 661 RHRNQLQMGKEPAPGAKKSNKPNQOTSKRHSDTFPEELKLTNMGSTKCSNTEKE 720  
QY 721 FVNPSPJREEKEKLEIVKVSNAEDPKDMLSGERVLOTERSVSSSISLPGDYGTQ 780  
DB 721 FVNPSPJREEKEKLEIVKVSNAEDPKDMLSGERVLOTERSVSSSISLPGDYGTQ 780  
QY 781 ESISLLEVSTLGAKAKTEPNKCVSQCAFENPKGLIHGSKDNDRNDTEGFKYPLGHEVMS 840  
DB 781 ESISLLEVSTLGAKAKTEPNKCVSQCAFENPKGLIHGSKDNDRNDTEGFKYPLGHEVMS 840  
QY 841 RETSIEMEESELDQYLOMTFVKRSKRSFALFSNPGNAEBCATPSAHSGLKQSPRYT 900  
DB 841 RETSIEMEESELDQYLOMTFVKRSKRSFALFSNPGNAEBCATPSAHSGLKQSPRYT 900  
QY 901 FECEQKEENQGNKNEINIKPVQITVITAGPVVGOKDPVDNAKCSIKGSRCFLSSQPRG 960  
DB 901 FECEQKEENQGNKNEINIKPVQITVITAGPVVGOKDPVDNAKCSIKGSRCFLSSQPRG 960  
QY 961 NETGLITPNKHGLQNPYRIPLEPIKSFVKTKCKKNLLENPEEHSMSPEREMENIP 1020  
DB 961 NETGLITPNKHGLQNPYRIPLEPIKSFVKTKCKKNLLENPEEHSMSPEREMENIP 1020  
QY 1021 STVSTISRRNIRENVKREASSNINEVSGSTNEVGSINIEISSPENIOAELGRNRGKL 1080  
DB 1021 STVSTISRRNIRENVKREASSNINEVSGSTNEVGSINIEISSPENIOAELGRNRGKL 1080  
QY 1081 NAMLRGLVLOPEVYKOSLPGSNCKHPEIKKQEEVEVYQVTNDFSPYLLSDNLEQPMSS 1140  
DB 1081 NAMLRGLVLOPEVYKOSLPGSNCKHPEIKKQEEVEVYQVTNDFSPYLLSDNLEQPMSS 1140  
QY 1141 HASOVCSFPPDLDLDGGEIKEDTSFAENDIKSSAVFSKSVQKGLSRSPPFTHTLQ 1200  
DB 1141 HASOVCSFPPDLDLDGGEIKEDTSFAENDIKSSAVFSKSVQKGLSRSPPFTHTLQ 1200  
QY 1201 GYRGAKKLEESSEENLSEDEELPCQHLIFKVVNIIPESQSRHSTVATECLSKTTEENL 1260  
DB 1201 GYRGAKKLEESSEENLSEDEELPCQHLIFKVVNIIPESQSRHSTVATECLSKTTEENL 1260  
QY 1261 LSLKNSLNDCSNOVILAKASQEHHLSEFTKCSASLFSQCSLELDLJANTNTQDPFLIGS 1320  
DB 1261 LSLKNSLNDCSNOVILAKASQEHHLSEFTKCSASLFSQCSLELDLJANTNTQDPFLIGS 1320  
QY 1321 SKQMRHQSOGVGLSDKELVSDDERGTCGLEENNOEDQSMNSNLGEAASGCESTVSSE 1380  
DB 1321 SKQMRHQSOGVGLSDKELVSDDERGTCGLEENNOEDQSMNSNLGEAASGCESTVSSE 1380  
QY 1381 DCSGLSSQSDILFTQORDIMOHLIKLQOEMAELEAVLEOHOSOPSNSYPSIISSSALE 1440  
DB 1381 DCSGLSSQSDILFTQORDIMOHLIKLQOEMAELEAVLEOHOSOPSNSYPSIISSSALE 1440  
QY 1441 DLRNPDOSTEKAVALTSQKSESEYPIQNPBGCSADKFEVSADOSTSKNKEPGEVRSFSPK 1500  
DB 1441 DLRNPDOSTEKAVALTSQKSESEYPIQNPBGCSADKFEVSADOSTSKNKEPGEVRSFSPK 1500  
QY 1501 CPSLDRWYMHSGSGSLQNRNTPSQQPELLIKYVDVEEQOLEEGPHDLTETSTLPRQDLEG 1560  
DB 1501 CPSLDRWYMHSGSGSLQNRNTPSQQPELLIKYVDVEEQOLEEGPHDLTETSTLPRQDLEG 1560  
QY 1561 TPYLESGISLFSDDPDSPDEPRAPEASAVGNIIPSTSTALKVPQLKVAESQSPAAAHTT 1620  
DB 1561 TPYLESGISLFSDDPDSPDEPRAPEASAVGNIIPSTSTALKVPQLKVAESQSPAAAHTT 1620  
QY 1621 DTAGYNAMEESVSRKPELJASTERVKNRSMNVVSGLTPPEEFMLVYKPARRHHITLNL 1680



```

Db      1621 DTAGNAMESVSSEKPELTAETERNKRMVWVSGLTPEEFMLVYKFAKRNHITLTNLI 1660
Qy      1681 TEETHVVMKTDAFVCEERTLKTYFLGAGKMWVSYFWMTOSIKERKMLNDHEVEGVGY 1740
Db      1681 TEETHVVMKTDAFVCEERTLKTYFLGAGKMWVSYFWMTOSIKERKMLNDHEVEGVGY 1740
Qy      1741 VNGRNHGGPKRARESDQRKIFRGLEICCYGPTNMPDQLEMMVQLGASVVKELSSFTL 1800
Db      1741 VNGRNHGGPKRARESDQRKIFRGLEICCYGPTNMPDQLEMMVQLGASVVKELSSFTL 1800
Qy      1801 GTGCHPIVYVQPDAMTENDNGFHAIGOMCEAPVYTVRENVLDVSVALYQCOELDTYLIPIPIH 1860
Db      1801 GTGCHPIVYVQPDAMTENDNGFHAIGOMCEAPVYTVRENVLDVSVALYQCOELDTYLIPIPIH 1860
Qy      1861 SHY 1863
Db      1861 SHY 1863
Qy      1861 SHY 1863
Db      1861 SHY 1863

RESULT 5
AAR81481
ID AAR81481 standard; Protein; 1863 AA.
AC AAR81481;
DE 30-SEP-1996 (first entry)
KW BRCA1.
XX Cancer therapy: breast and ovarian cancer predisposing gene; immunogen;
XX antibody production; germline alteration; probe; lesion neoplasia; human;
XX gene therapy: protein replacement therapy; protein mimetic; BRCA1.
OS Homo sapiens.
XX W09605306-A2.
XX 22-FEB-1996.
XX 11-AUG-1995; 95WO-US10202.
XX 07-JUN-1995; 95US-0483553.
XX 12-AUG-1994; 94US-0289221.
XX 02-SEP-1994; 94US-0300266.
XX 16-SEP-1994; 94US-0308104.
XX 29-NOV-1994; 94US-0348824.
XX 24-MAR-1995; 95US-0409305.
XX 07-JUN-1995; 95US-0480784.
XX (MYRI-) MYRIAD GENETICS INC.
XX (CANC-) CANCER INST.
XX (RECH-) CENT RECH DU CHUL.
XX Durocher F, Emi M, Nakamura Y, Simard J, Shattuck-Eidens DM;
XX WPI; 1996-139702/14.
XX N-PSDB; AAT17438.
XX
XX New nucleic acid and polypeptide for mutant or polymorphic BRCA1
XX gene - for diagnosis and therapy of human breast and ovarian cancer
XX and for diagnosing pre-disposition to these cancers
XX
XX Claim 1; Page 119-128; 218pp; English.
XX
XX This sequence represents the protein encoded by the human breast and
XX ovarian cancer predisposing gene (BRCA1). Mutations of this sequence
XX (see AAR81483-R81497 and AAR81499-R81546) can be used as immunogens for
XX antibody production. The mutant BRCA1 genes have at least 1 mutation or
XX polymorphism in comparison to the cDNA encoding this sequence. By
XX detecting a germline alteration in this gene, a predisposition for
XX breast and ovarian cancer can be diagnosed. In one method, BRCA1 mRNA
XX isolated from a tissue sample from a subject has a probe, corresponding

```

CC to a fragment of the cDNA encoding this sequence (or an allele-specific  
 CC probe for a mutation of it sequence), added to it. The conditions allow  
 CC for hybridisation of the probe to the mRNA, and any hybridisation which  
 CC occurs is detected. Alternatively the BRCA1 gene in the tissue sample  
 CC is isolated, and a shift in electrophoretic mobility of single stranded  
 CC DNA from the sample on a non-denaturing polyacrylamide gel indicates a  
 CC mutation. These methods of detection can also diagnose a lesion  
 CC neoplasia associated with the BRCA1 locus. The methods may be used in  
 CC gene therapy, protein replacement therapy and protein mimetics, and may  
 CC be used to screen for drugs in cancer therapy.

XX Sequence 1863 AA:

Query: Match 99.9%; Score 9642; DB 17; Length 1863;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1862; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy      1 MDLSALRVEEYQVYINAMOKITLCEPTCLELKEPVSTKCDHIFCKFCKMLKLNCKKPSQ 60
Db      1 MDLSALRVEEYQVYINAMOKITLCEPTCLELKEPVSTKCDHIFCKFCKMLKLNCKKPSQ 60
Qy      61 CPLCKNDITKRSLOESTRFSQVLYEELKTIICAPOLDTGLEANSYNFAKKENNSPEHLKD 120
Db      61 CPLCKNDITKRSLOESTRFSQVLYEELKTIICAPOLDTGLEANSYNFAKKENNSPEHLKD 120
Qy      121 EVSIIQSMGYNRNRAKRLQSEPNPSLOETSLSYQLSNLGTVRTLRKRIQIQPKTSYVI 180
Db      121 EVSIIQSMGYNRNRAKRLQSEPNPSLOETSLSYQLSNLGTVRTLRKRIQIQPKTSYVI 180
Qy      181 ELGSDSEEDIVNKATYCSVGDQLLQITPOGTRDEISLDSAKKACERSETVNTTEHHQ 240
Db      181 ELGSDSEEDIVNKATYCSVGDQLLQITPOGTRDEISLDSAKKACERSETVNTTEHHQ 240
Qy      181 ELGSDSEEDIVNKATYCSVGDQLLQITPOGTRDEISLDSAKKACERSETVNTTEHHQ 240
Db      181 ELGSDSEEDIVNKATYCSVGDQLLQITPOGTRDEISLDSAKKACERSETVNTTEHHQ 240
Qy      241 PSNNDLMTTEKRAERHPEKYYQSSVNLHVEPCGTTHASSLOHENSLLTKDRNVE 300
Db      241 PSNNDLMTTEKRAERHPEKYYQSSVNLHVEPCGTTHASSLOHENSLLTKDRNVE 300
Qy      301 KAEFCNKSQPGILARSOHNWAGSEKTCNDRTPTSTCKKYDLNADPLCERKENNKKLPC 360
Db      301 KAEFCNKSQPGILARSOHNWAGSEKTCNDRTPTSTCKKYDLNADPLCERKENNKKLPC 360
Qy      361 SENPDTEDEVPIITLANSIOKVNEMFSRSDLLGSDSDHGESESNKAVADVLDVLEVD 420
Db      361 SENPDTEDEVPIITLANSIOKVNEMFSRSDLLGSDSDHGESESNKAVADVLDVLEVD 420
Qy      421 EYSSSEKIDLLASDPHEALICKSERHKSVESENIDDKIFGKTYRKKASLPNLSHTEN 480
Db      421 EYSSSEKIDLLASDPHEALICKSERHKSVESENIDDKIFGKTYRKKASLPNLSHTEN 480
Qy      481 LIIGAFVTEPOIIQERPLTNKLRKRRPTSGILHPEDEIRKADLAVOKTPKINOGTNOTE 540
Db      481 LIIGAFVTEPOIIQERPLTNKLRKRRPTSGILHPEDEIRKADLAVOKTPKINOGTNOTE 540
Qy      541 QNGOVNMTNSGHENKTKGDSIONEKNPNPIESLEKESAFTYAEPITSSISNMLELNI 600
Db      541 QNGOVNMTNSGHENKTKGDSIONEKNPNPIESLEKESAFTYAEPITSSISNMLELNI 600
Qy      601 HNSKAPKKNRLRKRSSRTHALELVYSRNLSPNCTEIQIDSCSSSEELIKKKYNNMPV 660
Db      601 HNSKAPKKNRLRKRSSRTHALELVYSRNLSPNCTEIQIDSCSSSEELIKKKYNNMPV 660
Qy      661 RHRSRNLQLMGKPEPATGAKKSNKPNEDOTSKRHSDPTPELKLINAPSPFKCSNTSLKE 720
Db      661 RHRSRNLQLMGKPEPATGAKKSNKPNEDOTSKRHSDPTPELKLINAPSPFKCSNTSLKE 720
Qy      721 FVNPSPLEREKEKLETVKVSNNAEDEPKDMLSGERVLTQERSVSSSISLVPGTDVGTQ 780
Db      721 FVNPSPLEREKEKLETVKVSNNAEDEPKDMLSGERVLTQERSVSSSISLVPGTDVGTQ 780
Qy      781 EESISLEVSTLGAKEIKPKKVCVQCAAFENPKGLIHGCSKDNNDDEGFRPYLGHEVNH 840
Db      781 EESISLEVSTLGAKEIKPKKVCVQCAAFENPKGLIHGCSKDNNDDEGFRPYLGHEVNH 840

```

```

QY 841 RETSIEMESELDAOYLQNTFKVSKRSQFALFSPNGAEECATFSAHSGSLKKQSPKV 900
  |||||
Db 841 RETSIEMESELDAOYLQNTFKVSKRSQFALFSPNGAEECATFSAHSGSLKKQSPKV 900
QY 901 FECEQKEENQKKNESINIKPVQTVNITAGFVVGOKKPVNNAKCSIGGSRFCLSSQFRG 960
  |||||
Db 901 FECEQKEENQKKNESINIKPVQTVNITAGFVVGOKKPVNNAKCSIGGSRFCLSSQFRG 960
QY 961 NETGLTPNKHGLLOPPIRPLPIKSFYTKCKKLLLENFEHSMSPEREKGNINIP 1020
  |||||
Db 961 NETGLTPNKHGLLOPPIRPLPIKSFYTKCKKLLLENFEHSMSPEREKGNINIP 1020
QY 961 NETGLTPNKHGLLOPPIRPLPIKSFYTKCKKLLLENFEHSMSPEREKGNINIP 1020
  |||||
Db 961 NETGLTPNKHGLLOPPIRPLPIKSFYTKCKKLLLENFEHSMSPEREKGNINIP 1020
QY 1021 STVSTISRNINIRENVEKEASSNINEVSGSTNEVGSSINEIGSSDENIQALGNRGPKL 1080
  |||||
Db 1021 STVSTISRNINIRENVEKEASSNINEVSGSTNEVGSSINEIGSSDENIQALGNRGPKL 1080
QY 1081 NAMLRIGVLOPEYKOSLPGSNCKHPEIKOEYEVQTVNTDPSPLIDNLEQPMGSS 1140
  |||||
Db 1081 NAMLRIGVLOPEYKOSLPGSNCKHPEIKOEYEVQTVNTDPSPLIDNLEQPMGSS 1140
QY 1081 NAMLRIGVLOPEYKOSLPGSNCKHPEIKOEYEVQTVNTDPSPLIDNLEQPMGSS 1140
  |||||
Db 1081 NAMLRIGVLOPEYKOSLPGSNCKHPEIKOEYEVQTVNTDPSPLIDNLEQPMGSS 1140
QY 1141 HASQVSETPDDLDDGETKEOTSFPAENDIKESSAVFSKSVQSGELSRSPFTHTHIAQ 1200
  |||||
Db 1141 HASQVSETPDDLDDGETKEOTSFPAENDIKESSAVFSKSVQSGELSRSPFTHTHIAQ 1200
QY 1201 GYRGAKKLESSEENLSEDELLPCFOHLLFGVNNIPSGSTRHSTVATECLSKNTEBL 1260
  |||||
Db 1201 GYRGAKKLESSEENLSEDELLPCFOHLLFGVNNIPSGSTRHSTVATECLSKNTEBL 1260
QY 1261 LSLKNSLNCOSNYIILAKAQEHHLSEPTKCSASLPSQSCSELEDLTANNTQDPLIGS 1320
  |||||
Db 1261 LSLKNSLNCOSNYIILAKAQEHHLSEPTKCSASLPSQSCSELEDLTANNTQDPLIGS 1320
QY 1321 SKQMRQSESOQVGLSDKELVSDDEERGTLLEENQEOQMDSNLGAASCESETVSE 1380
  |||||
Db 1321 SKQMRQSESOQVGLSDKELVSDDEERGTLLEENQEOQMDSNLGAASCESETVSE 1380
QY 1381 DCSGSLSSQSDILTTQQRDTMQHNLKIQOEMAELEAVLEQHGQSPNSYPSIISDSALE 1440
  |||||
Db 1381 DCSGSLSSQSDILTTQQRDTMQHNLKIQOEMAELEAVLEQHGQSPNSYPSIISDSALE 1440
QY 1441 DLRNPEOSTSEKAVLTSQKSEYPISONPEGLSADKFEVADSDSTSKNKEPVERSSPSK 1500
  |||||
Db 1441 DLRNPEOSTSEKAVLTSQKSEYPISONPEGLSADKFEVADSDSTSKNKEPVERSSPSK 1500
QY 1501 CPSLDDRMWYHSCSGSLQNRNPSQBELIKVVDVEEQOLEESGPHDLTFEVSYPRODLEG 1560
  |||||
Db 1501 CPSLDDRMWYHSCSGSLQNRNPSQBELIKVVDVEEQOLEESGPHDLTFEVSYPRODLEG 1560
QY 1561 TPYLESIGLFSDDPSPSEDRAPEARSAGNIPSSTSALKVPQLKVAESAQSPAAAHNT 1620
  |||||
Db 1561 TPYLESIGLFSDDPSPSEDRAPEARSAGNIPSSTSALKVPQLKVAESAQSPAAAHNT 1620
QY 1621 DRAGYAAEESYSREKPELTASTERYNKRMSVVSGLTPEEFMLVYKFAKHHITLNL 1680
  |||||
Db 1621 DRAGYAAEESYSREKPELTASTERYNKRMSVVSGLTPEEFMLVYKFAKHHITLNL 1680
QY 1681 TEETTHVAKTAEFEYCERLTLYLGLIAGKMWVSFWVYOSTIKERKMLNEHDFEYAGDV 1740
  |||||
Db 1681 TEETTHVAKTAEFEYCERLTLYLGLIAGKMWVSFWVYOSTIKERKMLNEHDFEYAGDV 1740
QY 1741 VNGRNHQGKRARESODRKIFRGLEICCGPFTNMPDLEMMVOLCGASVKELESFTL 1800
  |||||
Db 1741 VNGRNHQGKRARESODRKIFRGLEICCGPFTNMPDLEMMVOLCGASVKELESFTL 1800
QY 1801 GTGVHPRIYVOPDAMVEDNGFHAIGOMCEAPVYTRREMYLDVALXOCQELDTYLIPQIPH 1860
  |||||
Db 1801 GTGVHPRIYVOPDAMVEDNGFHAIGOMCEAPVYTRREMYLDVALXOCQELDTYLIPQIPH 1860
QY 1861 SHY 1863
  |||
Db 1861 SHY 1863

```

```

RESULT 6
ID AAR91208
XX AAR91208 standard; Protein: 1863 AA.
AC AAR91208;
XX
DT 05-JUN-1996 (first entry)
XX
DE BRCA1..breast and ovarian cancer predisposing gene product.
XX
KW BRCA1..breast cancer; ovary cancer; predisposing gene;
KW susceptible/idiolity gene; protein replacement therapy; diagnosis;
KW prognosis.
XX
OS Homo sapiens.
XX
PN W09605307-A2.
PD 22-FEB-1996.
XX
PE 11-AUG-1995; 95WO-US10203.
PR 07-JUN-1995; 95US-0488011.
PR 12-AUG-1994; 94US-0289221.
PR 02-SEP-1994; 94US-0300266.
PR 16-SEP-1994; 94US-0308104.
PR 29-NOV-1994; 94US-0348824.
PR 24-MAR-1995; 95US-0409305.
PR 07-JUN-1995; 95US-0483554.
PR 07-JUN-1995; 95US-0487002.
XX
PA (MYRI-) MYRIAD GENETICS INC.
PA (UTAH ) UNIV UTAH RES FOUND.
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Eutreal AP, Goldgar DE, Harsiman KD, Kamb A, Miki Y;
PI Shattuck-Eidens DM, Skolnick MH, Swenson J, Tavtigian SV;
PI Wiseman RW;
XX
DR WPI: 1996-139703/14.
DR N-PSDB: AAT18310.
XX
PT New isolated human cancer predisposing gene, BRCA1 - used to develop
PT prods. for diagnosis, prognosis and therapy of cancers, partic.
PT breast and ovarian cancers
XX
PS Claim 1; Page 117-124; 190pp; English.
XX
CC The BRCA1 polypeptide (AAR91208) is the product of the human BRCA1
CC gene (AAT18310). It can be obtd. by insertion of BRCA1 DNA into a
CC vector and expression in host cells; products of mutated BRCA1
CC genes, associated in humans with a predisposition to breast and
CC ovarian cancer, can also be produced. BRCA1 polypeptides are used
CC as immunogens for the prodn. of antibodies, or to screen therapeutic
CC drugs. They can also be used in protein replacement therapy.
XX
SQ Sequence 1863 AA.
Query Match 99.9%; Score 9642; DB 17; Length 1863;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1862; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MDLSALRVEVQVIVAMOKITLCPICLIEKEPVSTKCDHIECKFCMKLILNOKKGPSQ 60
  |||||
Db 1 MDLSALRVEVQVIVAMOKITLCPICLIEKEPVSTKCDHIECKFCMKLILNOKKGPSQ 60
QY 61 CPLCKNDITKRSLOESTRESQVLEELKTTICAFOLDTGLEVANSYNFAKKNNSPEHLKD 120
  |||||
Db 61 CPLCKNDITKRSLOESTRESQVLEELKTTICAFOLDTGLEVANSYNFAKKNNSPEHLKD 120
QY 61 CPLCKNDITKRSLOESTRESQVLEELKTTICAFOLDTGLEVANSYNFAKKNNSPEHLKD 120
  |||||
Db 61 CPLCKNDITKRSLOESTRESQVLEELKTTICAFOLDTGLEVANSYNFAKKNNSPEHLKD 120
QY 121 EVSIIOSMGYRNRAKRLLOSPEPNPSLOETSLISVOLSNLGTIVRLTRKORIOPOKTSYVI 180
  |||||
Db 121 EVSIIOSMGYRNRAKRLLOSPEPNPSLOETSLISVOLSNLGTIVRLTRKORIOPOKTSYVI 180

```

QY 181 ELGSDSEPTVNAKATYCSVDOELLQITPOGTRDEISLDSAKKACEFSETDVTNTEHHO 240  
 DB 181 ELGSDSEPTVNAKATYCSVDOELLQITPOGTRDEISLDSAKKACEFSETDVTNTEHHO 240  
 QY 241 PSNNIDNTTEKRAAEHHPKYOSSVSNLHVEPCGNTTHASSLOHENSLLLTKDRMAYE 300  
 DB 241 PSNNIDNTTEKRAAEHHPKYOSSVSNLHVEPCGNTTHASSLOHENSLLLTKDRMAYE 300  
 QY 301 KAFCCKSKQPGIARQOHNRWAGSKETCDNRTPSTEEKYVDLADLCEKRENNKQKLC 360  
 DB 301 KAFCCKSKQPGIARQOHNRWAGSKETCDNRTPSTEEKYVDLADLCEKRENNKQKLC 360  
 QY 361 SENRPTDEVPMTLLNSSIQKVENMFSRDELLGSDSHDGESESNAKADVLDVNEVD 420  
 DB 361 SENRPTDEVPMTLLNSSIQKVENMFSRDELLGSDSHDGESESNAKADVLDVNEVD 420  
 QY 421 EYSGSSEKIDLLASDPHEALICKSERVHSKSVESNIEDKIFGKYRKASLPNLSHVTEN 480  
 DB 421 EYSGSSEKIDLLASDPHEALICKSERVHSKSVESNIEDKIFGKYRKASLPNLSHVTEN 480  
 QY 481 LIIGAVTEPQIIQERPLNKLKRRRPTSGLHPEDFIKADLAVQKTEPMINOGNOTE 540  
 DB 481 LIIGAVTEPQIIQERPLNKLKRRRPTSGLHPEDFIKADLAVQKTEPMINOGNOTE 540  
 QY 541 ONGOVNITNSGHEKTKGDSIONEKNPPIESLEKESAFKTKAEPJISSISMELELNI 600  
 DB 541 ONGOVNITNSGHEKTKGDSIONEKNPPIESLEKESAFKTKAEPJISSISMELELNI 600  
 QY 601 HNSKAPKKNRLRRKSTRIHALELVVSNLSPPNCTELOIDSCSSSEBIKKKKYQOMV 660  
 DB 601 HNSKAPKKNRLRRKSTRIHALELVVSNLSPPNCTELOIDSCSSSEBIKKKKYQOMV 660  
 QY 661 RHSRNQIOMEKREPATGAKSKNPNQOTSKRHSODTFPELKLITNAGSTKCSNTSELE 720  
 DB 661 RHSRNQIOMEKREPATGAKSKNPNQOTSKRHSODTFPELKLITNAGSTKCSNTSELE 720  
 QY 721 FVNPSLPREKKEKLETTVAVSNNAEDPKDMLSGERVLOTERSVESSISLVPGTDTYGO 780  
 DB 721 FVNPSLPREKKEKLETTVAVSNNAEDPKDMLSGERVLOTERSVESSISLVPGTDTYGO 780  
 QY 781 ESTISLEVTSLGAKTEPKKCVSOCAAFENPKGLIHGCSKDNBNDETEGFKYPLGHEVNS 840  
 DB 781 ESTISLEVTSLGAKTEPKKCVSOCAAFENPKGLIHGCSKDNBNDETEGFKYPLGHEVNS 840  
 QY 841 RETSIEMESELDAOYLONTFKVSKROSFALFSPNPNABEBCATPSAHSGLKKOSPXYT 900  
 DB 841 RETSIEMESELDAOYLONTFKVSKROSFALFSPNPNABEBCATPSAHSGLKKOSPXYT 900  
 QY 901 FECEQKEENOGKNESNIKPVQTVNITAGFPVVGQKDPVDNAKCSIKGSRFCLSSQFRG 960  
 DB 901 FECEQKEENOGKNESNIKPVQTVNITAGFPVVGQKDPVDNAKCSIKGSRFCLSSQFRG 960  
 QY 961 NETGLITPKHGLOLPYRIPPLPIKSFVTKCKKNLLEENFEHSMSPEREMENIP 1020  
 DB 961 NETGLITPKHGLOLPYRIPPLPIKSFVTKCKKNLLEENFEHSMSPEREMENIP 1020  
 QY 1021 STVSTISRNINIRENFKEASSNININYGSTNINIEGSSIDENIOAELGRNGPKL 1080  
 DB 1021 STVSTISRNINIRENFKEASSNININYGSTNINIEGSSIDENIOAELGRNGPKL 1080  
 QY 1081 NAMLRIGVQPEPVYKQSLPGSNCKHPEIKQEEVEEVQTVNDFSYLLTSDMLEOPMGS 1140  
 DB 1081 NAMLRIGVQPEPVYKQSLPGSNCKHPEIKQEEVEEVQTVNDFSYLLTSDMLEOPMGS 1140  
 QY 1141 HASOVCESETPDDLDDGEIKEDTSPFAENDIKESSAFAFSKVOGELSRSPFTHTLQ 1200  
 DB 1141 HASOVCESETPDDLDDGEIKEDTSPFAENDIKESSAFAFSKVOGELSRSPFTHTLQ 1200  
 QY 1201 GYRGAKKLESSEENLSEDELPCHOLFGKVVNNIPSOSTRHSVATECLSKNTEENL 1260  
 DB 1201 GYRGAKKLESSEENLSEDELPCHOLFGKVVNNIPSOSTRHSVATECLSKNTEENL 1260

QY 1261 LSLKNSLNDSCSNVILAKASOEHHLSSEETKCSALFSSQCSLELDTANTNDPPLIGS 1320  
 DB 1261 LSLKNSLNDSCSNVILAKASOEHHLSSEETKCSALFSSQCSLELDTANTNDPPLIGS 1320  
 QY 1321 SKOMRHQSESQGVGLSDKELVSDDEBGTGLEENNOEOMSNDNLGEAASCESETSYSE 1380  
 DB 1321 SKOMRHQSESQGVGLSDKELVSDDEBGTGLEENNOEOMSNDNLGEAASCESETSYSE 1380  
 QY 1381 DCSGLSSQSDILITTOOROTMOMHLIKLOOEMALEVLBDHOGQSPNSPSTISDSALE 1440  
 DB 1381 DCSGLSSQSDILITTOOROTMOMHLIKLOOEMALEVLBDHOGQSPNSPSTISDSALE 1440  
 QY 1441 DLRNPEOSTSEKAVLTSOKSSEYPIQONPEGLSADFEVSADSTSKNKEPVERSSPSK 1500  
 DB 1441 DLRNPEOSTSEKAVLTSOKSSEYPIQONPEGLSADFEVSADSTSKNKEPVERSSPSK 1500  
 QY 1501 CPSLDRWYMHSCSGSLQNRNYPQOELLIKVVDVEEQOLEESGPHDLTETSYLPQDLEG 1560  
 DB 1501 CPSLDRWYMHSCSGSLQNRNYPQOELLIKVVDVEEQOLEESGPHDLTETSYLPQDLEG 1560  
 QY 1561 TPYLESGISLFDSDPESDEDPAPESARVNTPSSTSALKVPQLKVAESAQSPAANT 1620  
 DB 1561 TPYLESGISLFDSDPESDEDPAPESARVNTPSSTSALKVPQLKVAESAQSPAANT 1620  
 QY 1621 DTAGYNAMEESVSREKPELTASTERYNKRMSWVSGLTPEEPMLYKFKARKHITLTNLI 1680  
 DB 1621 DTAGYNAMEESVSREKPELTASTERYNKRMSWVSGLTPEEPMLYKFKARKHITLTNLI 1680  
 QY 1681 TEETHVVKKTDAEFYCERLTAKYFLIGIAGKVVVSFYFWYOSIKERKMLENHDFEVRGCV 1740  
 DB 1681 TEETHVVKKTDAEFYCERLTAKYFLIGIAGKVVVSFYFWYOSIKERKMLENHDFEVRGCV 1740  
 QY 1741 VNCRNHQGRKRAESQDRKIFRGLETCCYGPFTNMPTDOLLEWVQLCGASVKESSFTL 1800  
 DB 1741 VNCRNHQGRKRAESQDRKIFRGLETCCYGPFTNMPTDOLLEWVQLCGASVKESSFTL 1800  
 QY 1801 GTGVHPYVVPDPAWEDNGFHAIGOMCEAPVYTRFVMDSVVALYQCELDYVLLIQIPH 1860  
 DB 1801 GTGVHPYVVPDPAWEDNGFHAIGOMCEAPVYTRFVMDSVVALYQCELDYVLLIQIPH 1860  
 QY 1861 SHY 1863  
 DB 1861 SHY 1863  
 RESULT 7  
 AAY32033  
 ID AAY32033 standard; protein; 1863 AA.  
 XX  
 AC AAY32033;  
 XX  
 DT 05-JAN-2000 (first entry)  
 XX  
 DE Human BRCA1 protein.  
 XX  
 KW BRCA1: p53 protein; p21 gene; human; tumour suppressor;  
 KW transcriptional activator; breast cancer; cell proliferation;  
 KW apoptosis; diagnosis; anticancer; antitumour; drug screening.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key location/Qualifiers  
 FT Binding-site 224..500  
 FT Region /note="binding site of p53 protein"  
 FT 499..510  
 PN /note="NLS region"  
 XX  
 PD WO9950280-A1.  
 XX  
 PF 07-OCT-1999.  
 XX  
 PF 31-MAR-1999; 99MO-0507150.  
 XX

PR 31-MAR-1998; 98US-0080146.  
XX (UPE-) UNIV PENNSYLVANIA.  
XX El-Deliry WS, Weber BL;  
PI WPI, 1999-601319/51.  
XX  
XX Nucleic acid involved in BRCA-1-mediated control of transcriptional  
PT regulation of tumour suppressor genes and related peptides, used to  
PT screen for modulators for use as anticancer agents  
PS  
XX Claim 9: Page 76-80; 94pp; English.  
XX  
XX This sequence represents the human BRCA1 protein. BRCA-1  
CC inhibits S-phase cell-cycle progression by transactivating  
CC expression of p21 in a p53-independent manner. It also  
CC interacts with p53 (see AAY32034) in vitro and in vivo, and  
CC increases p53-dependent transcription from the p21 (see AA220131)  
CC and bax promoters. These newly discovered pathways of BRCA1 action  
CC provide novel targets to which pharmaceutical agents capable of  
CC affecting cell proliferation and apoptosis can be isolated, and  
CC can then be used in the treatment and control of cellular  
CC proliferation disorders, such as breast cancer. Assays and  
CC compositions for identifying compounds that enhance or repress  
CC cellular proliferation via these BRCA1-mediated pathways are  
CC disclosed.  
CC  
XX  
SQ Sequence 1863 AA;  
Query Match 99.9%; Score 9642; DB 20; Length 1863;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1862; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MDLSALRYEYQNVINAMOKILECPICLELKEPEYTKCDHIFCKFCMLKLNOKKPSQ 60  
DB 1 MDLSALRYEYQNVINAMOKILECPICLELKEPEYTKCDHIFCKFCMLKLNOKKPSQ 60  
QY 61 CPLCKNDITKNSLOESTRETSQVLEELIKTICAFOLDTGEVANSYNFAKKENNSPEHLD 120  
DB 61 CPLCKNDITKNSLOESTRETSQVLEELIKTICAFOLDTGEVANSYNFAKKENNSPEHLD 120  
QY 121 EVSTIQQSGYRNRARLLQSEPEPNSLOTSLSVOLSNLGYRTLTQRIOPOKTSYVI 180  
DB 121 EVSTIQQSGYRNRARLLQSEPEPNSLOTSLSVOLSNLGYRTLTQRIOPOKTSYVI 180  
QY 181 ELGSDSEDTYKATYCVSGDOELQITPOGTRDEISLDSAKKACEFSETVNTNEHQ 240  
DB 181 ELGSDSEDTYKATYCVSGDOELQITPOGTRDEISLDSAKKACEFSETVNTNEHQ 240  
QY 241 PSNNDLNTTEKRAERHPEKYOGSSVNLHVEPCGTNTHTASSLOHENSLLTKDRMVE 300  
DB 241 PSNNDLNTTEKRAERHPEKYOGSSVNLHVEPCGTNTHTASSLOHENSLLTKDRMVE 300  
QY 301 KAFCNKSQKOPGLARSOHNRAGSKETCNDRTPESTKRVYDNLADPLCEKRMNOKLPC 360  
DB 301 KAFCNKSQKOPGLARSOHNRAGSKETCNDRTPESTKRVYDNLADPLCEKRMNOKLPC 360  
QY 361 SENRDEDTVPMTLNSSTIOKYNEMFSRSDDELGSDSHDGESESNAAVADLVLENYD 420  
DB 361 SENRDEDTVPMTLNSSTIOKYNEMFSRSDDELGSDSHDGESESNAAVADLVLENYD 420  
QY 421 EYSSSSEKIDILASDPHEALICKSERVHSHSVESNIEEDIKFGTYRKKASLPNLSHVEN 480  
DB 421 EYSSSSEKIDILASDPHEALICKSERVHSHSVESNIEEDIKFGTYRKKASLPNLSHVEN 480  
QY 481 LIIAFYTEPQIOIERPLTNKLRKRRPTSGLHPEDFIKKADLAQKTPPEMNOCTNOTE 540  
DB 481 LIIAFYTEPQIOIERPLTNKLRKRRPTSGLHPEDFIKKADLAQKTPPEMNOCTNOTE 540  
QY 541 QNGOVAMITNSGHNKTKGDSIONEKNPNPIESLEKESAFKTKAPISSISNMLELNI 600  
DB 541 QNGOVAMITNSGHNKTKGDSIONEKNPNPIESLEKESAFKTKAPISSISNMLELNI 600

QY 601 HNSKAPKKNLRRKSSSTRHIALELVYSNLSPPNCTELQIDSCSSSEETIKKKKNQMPV 660  
DB 601 HNSKAPKKNLRRKSSSTRHIALELVYSNLSPPNCTELQIDSCSSSEETIKKKKNQMPV 660  
QY 661 RHRNRQLMEGKDPATGAKKSNKPNQOTSKRHDSOTFPPLKLTNAPGFTKCSNTSEKE 720  
DB 661 RHRNRQLMEGKDPATGAKKSNKPNQOTSKRHDSOTFPPLKLTNAPGFTKCSNTSEKE 720  
QY 721 FVNPSPREBEKEKLTAVVSNNAEDPKMLSGERVLOTESVSSSTSLVPGDVTQ 780  
DB 721 FVNPSPREBEKEKLTAVVSNNAEDPKMLSGERVLOTESVSSSTSLVPGDVTQ 780  
QY 781 ESISLLEVSITLAKTEPNKCVSOGAEPENPGLIHGCSKDRNDETEGFKPLGHEVNS 840  
DB 781 ESISLLEVSITLAKTEPNKCVSOGAEPENPGLIHGCSKDRNDETEGFKPLGHEVNS 840  
QY 841 RETSIMESEIDAOYLQNTFVYSKROSPALFSPNGNAEECATFSAHSGSLKOSPKVT 900  
DB 841 RETSIMESEIDAOYLQNTFVYSKROSPALFSPNGNAEECATFSAHSGSLKOSPKVT 900  
QY 901 FECEQKEENQKNESIKRVOJVTAGPPVVGOKDPVDNAKCSIKGSRFLCSSOPRG 960  
DB 901 FECEQKEENQKNESIKRVOJVTAGPPVVGOKDPVDNAKCSIKGSRFLCSSOPRG 960  
QY 961 NETGLITPKHGLQNPYRIPULFPIKSFVKTKCKKNLLEENFEBSHSPREMGNEPI 1020  
DB 961 NETGLITPKHGLQNPYRIPULFPIKSFVKTKCKKNLLEENFEBSHSPREMGNEPI 1020  
QY 1021 STVSTISRNNTRENNFKKASSNINDEVSSNINEGSSDENITQALGRNCPKL 1080  
DB 1021 STVSTISRNNTRENNFKKASSNINDEVSSNINEGSSDENITQALGRNCPKL 1080  
QY 1081 NAMLRGLVQPEYKOSLPGSNCKHPEIKQOEYEVQVNTVDSPLYLSNLEOPMSS 1140  
DB 1081 NAMLRGLVQPEYKOSLPGSNCKHPEIKQOEYEVQVNTVDSPLYLSNLEOPMSS 1140  
QY 1141 HASQVSETPDDLDDGEIKEDTSFAENDIKESSAVFSKYQKGLSPSPFTHTLQ 1200  
DB 1141 HASQVSETPDDLDDGEIKEDTSFAENDIKESSAVFSKYQKGLSPSPFTHTLQ 1200  
QY 1201 GYRGAKKLESSEENLSEDELPFOHLTGKVNINPISOSTRSHVATBELSKNTEBNL 1260  
DB 1201 GYRGAKKLESSEENLSEDELPFOHLTGKVNINPISOSTRSHVATBELSKNTEBNL 1260  
QY 1261 LSLKNSLNDCSNOVILAKASQEHLSSEETKCSALFSSQSELEDLTANTMTQDPFLIGS 1320  
DB 1261 LSLKNSLNDCSNOVILAKASQEHLSSEETKCSALFSSQSELEDLTANTMTQDPFLIGS 1320  
QY 1321 SKOMRHQSESGVGLSDKELVSDDEBRTGLENNQEQSMDSNLGEAASCESETSYSE 1380  
DB 1321 SKOMRHQSESGVGLSDKELVSDDEBRTGLENNQEQSMDSNLGEAASCESETSYSE 1380  
QY 1381 DCSGLSSOSDILTTQORTMQHNLKLDQEMAELEAVLEQHGSPSNYSIISDSSALE 1440  
DB 1381 DCSGLSSOSDILTTQORTMQHNLKLDQEMAELEAVLEQHGSPSNYSIISDSSALE 1440  
QY 1441 DLRNEQSTSEKAVLTISQKSESEYPISONPPEGLSADKFEVVSADSTSKNKEGVRSSPSK 1500  
DB 1441 DLRNEQSTSEKAVLTISQKSESEYPISONPPEGLSADKFEVVSADSTSKNKEGVRSSPSK 1500  
QY 1501 CPSLDDRWYMHSCSGSLQNRNVPQOEELIKVVDVEOQLJESGPHDLTETSYLPROLEG 1560  
DB 1501 CPSLDDRWYMHSCSGSLQNRNVPQOEELIKVVDVEOQLJESGPHDLTETSYLPROLEG 1560  
QY 1561 TPYLESGISLSDDESPSEDRAPEASRVGNITSSSALKVPOPLKVAESQSPAANTT 1620  
DB 1561 TPYLESGISLSDDESPSEDRAPEASRVGNITSSSALKVPOPLKVAESQSPAANTT 1620  
QY 1621 DTAGYNAMEESVSRKPELTASTERVNKRMSVMVSGLTPEEFMLVYFAPKHHHTLTNLI 1680  
DB 1621 DTAGYNAMEESVSRKPELTASTERVNKRMSVMVSGLTPEEFMLVYFAPKHHHTLTNLI 1680

QY 1681 TEETHVAVKTDAEVFCERTLKLYFLGIAGKVVVSYFWVTSIKERKMLNEHDFEVRGDV 1740  
 |||||||  
 Db 1681 TEETHVAVKTDAEVFCERTLKLYFLGIAGKVVVSYFWVTSIKERKMLNEHDFEVRGDV 1740  
 QY 1741 VNGRNQGRKRAESDDKIFRGLETICCGPFTNMPTDLEWVQICGASVYKELSSFTL 1800  
 |||||||  
 Db 1741 VNGRNQGRKRAESDDKIFRGLETICCGPFTNMPTDLEWVQICGASVYKELSSFTL 1800  
 QY 1801 GTGVHPVIVVOPDAMVEDNGFHAIGOMCAPVYTRVMVDSVALYOCQELDTYLLIPQIH 1860  
 |||||||  
 Db 1801 GTGVHPVIVVOPDAMVEDNGFHAIGOMCAPVYTRVMVDSVALYOCQELDTYLLIPQIH 1860  
 QY 1861 SHY 1863  
 |||  
 Db 1861 SHY 1863  
 |||

RESULT 8  
 ABG01696  
 ID ABG01696 standard; Protein; 2353 AA.  
 XX  
 AC ABG01696:  
 XX  
 DT 13-FEB-2002 (first entry)  
 XX  
 DE Novel human diagnostic protein #1687.  
 XX  
 KW Human: chromosome mapping; gene mapping; gene therapy; forensic;  
 food supplement; medical imaging; diagnostic; genetic disorder.  
 XX  
 OS Homo sapiens.  
 OS  
 PN WC0200175067-A2.  
 XX  
 PD 11-OCT-2001.  
 PF  
 PF 30-MAR-2001; 2001WC-0S08631.  
 XX  
 PR 31-MAR-2000; 2000US-0540217.  
 PR 23-AUG-2000; 2000US-0649167.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 PI Drmanac RT, Liu C, Tang YT;  
 XX  
 DR WPI: 2001-639362/73.  
 DR N-PSDB: AAS65883.  
 XX  
 PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity  
 XX  
 PS Claim 20; SEQ ID No 32055; 103pp; English.

The invention relates to isolated polynucleotide (I) and  
 polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 and gene mapping, and in recombinant production of (II). The  
 polynucleotides are also used in diagnostics as expressed sequence tags  
 for identifying expressed genes. (I) is useful in gene therapy techniques  
 to restore normal activity of (II) or to treat disease states involving  
 (II). (II) is useful for generating antibodies against it, detecting or  
 quantitating a polypeptide in tissue, as molecular weight markers and as  
 a food supplement. (II) and its binding partners are useful in medical  
 imaging of sites expressing (II). (I) and (II) are useful for treating  
 disorders involving aberrant protein expression or biological activity.  
 The polypeptide and polynucleotide sequences have applications in  
 diagnostics, forensics, gene mapping, identification of mutations  
 responsible for genetic disorders or other traits to assess biodiversity  
 and to produce other types of data and products dependent on DNA and  
 amino acid sequences. ABG0010-ABG30377 represent novel human  
 diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX

XX Sequence 2353 AA:

Query Match 99.9%; Score 9642; DB 22; Length 2353;  
 Best local Similarity 99.9%; Pred. No. 0;  
 Matches 1862; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDLSALRVEEYONVINAMOKILECPICLEIKPEVSTKCDHIFCKFCMKLLNOKKGPQ 60  
 |||||||  
 Db 491 MDLSALRVEEYONVINAMOKILECPICLEIKPEVSTKCDHIFCKFCMKLLNOKKGPQ 550  
 QY 61 CPLCKNDITKRSLQESTRSQVLEELLKTTICAFQDPTGLEAYNSYPAKKENNSPHELD 120  
 |||||||  
 Db 551 CPLCKNDITKRSLQESTRSQVLEELLKTTICAFQDPTGLEAYNSYPAKKENNSPHELD 610  
 QY 121 EYSIIOSMGYRNRAKRLQSEPNPSLOETSLVSQLSNLTGVYTLTKQRIOPQKTSYVI 180  
 |||||||  
 Db 611 EYSIIOSMGYRNRAKRLQSEPNPSLOETSLVSQLSNLTGVYTLTKQRIOPQKTSYVI 670  
 QY 181 ELGSDSEEDTVNKATYCSVGDELLQITPOGTRDEISLDSAKKACEFSETDVTNHHQ 240  
 |||||||  
 Db 671 ELGSDSEEDTVNKATYCSVGDELLQITPOGTRDEISLDSAKKACEFSETDVTNHHQ 730  
 QY 241 PSNNDLNTTEKRAAEHPKRYOGSSVSNLHVEPCGNTTASSIQHNSSILLTKDMNVE 300  
 |||||||  
 Db 731 PSNNDLNTTEKRAAEHPKRYOGSSVSNLHVEPCGNTTASSIQHNSSILLTKDMNVE 790  
 QY 301 KAEFCNKSQOPGLARSOHNRWAGSKETCNDRTPTSTKKVYDLNADPLCEKKEVKKOLPC 360  
 |||||||  
 Db 791 KAEFCNKSQOPGLARSOHNRWAGSKETCNDRTPTSTKKVYDLNADPLCEKKEVKKOLPC 850  
 QY 361 SENPRDTEVPWITLNSIQKVENWFSRDELDGSDSDHGESESNAAVADVLVNEVD 420  
 |||||||  
 Db 851 SENPRDTEVPWITLNSIQKVENWFSRDELDGSDSDHGESESNAAVADVLVNEVD 910  
 QY 421 EYSGSSEKIDLASDHEALICKSEEVHSHKSVESNIEDKIFGTYKKKASLPVLSVTEN 480  
 |||||||  
 Db 911 EYSGSSEKIDLASDHEALICKSEEVHSHKSVESNIEDKIFGTYKKKASLPVLSVTEN 970  
 QY 481 LIIGAFVTEPQIIQOERPLTNKLRKRRTSGLHPEDFIKKADLAVQKTEPMINQNTQTE 540  
 |||||||  
 Db 971 LIIGAFVTEPQIIQOERPLTNKLRKRRTSGLHPEDFIKKADLAVQKTEPMINQNTQTE 1030  
 QY 541 QNGQVNNITNSGHENKTKGDSIQNEKNPNPIESLEKESAFKTKAEDYSSISNMELELN 600  
 |||||||  
 Db 1031 QNGQVNNITNSGHENKTKGDSIQNEKNPNPIESLEKESAFKTKAEDYSSISNMELELN 1090  
 QY 601 HNSKAPKKRRLKRRKSTRIHALELVVSRNLSPNCTELQIDSCSSSEIKKKRYQMVY 660  
 |||||||  
 Db 1091 HNSKAPKKRRLKRRKSTRIHALELVVSRNLSPNCTELQIDSCSSSEIKKKRYQMVY 1150  
 QY 661 RISRNLQMEGKEPATGAKSKNPNQOTSKRHSDFPELKLTUNAGSFTKCSNTSELKE 720  
 |||||||  
 Db 1151 RISRNLQMEGKEPATGAKSKNPNQOTSKRHSDFPELKLTUNAGSFTKCSNTSELKE 1210  
 QY 721 FVNPSPREKEKEKLETVKVSNNAEDEPKDMLSGERVLTQERSVESSISLVPGTDYGRQ 780  
 |||||||  
 Db 1211 FVNPSPREKEKEKLETVKVSNNAEDEPKDMLSGERVLTQERSVESSISLVPGTDYGRQ 1270  
 QY 781 ESTSLLEVSTLKAKEPKKCVSQCFAFNPGLIHGCSKDNNDNDEGRYPGLGHEVNS 840  
 |||||||  
 Db 1271 ESTSLLEVSTLKAKEPKKCVSQCFAFNPGLIHGCSKDNNDNDEGRYPGLGHEVNS 1330  
 QY 841 RETSIEMESEDADQTLQMTFVYSKROSFALESNPNAGSEECATFAHSGSLKQSPKTY 900  
 |||||||  
 Db 1331 RETSIEMESEDADQTLQMTFVYSKROSFALESNPNAGSEECATFAHSGSLKQSPKTY 1390  
 QY 901 FECEOKEENQGNESNIKRVQTVNITAGFPVVGQDKPVDNAKCSIKGSSRFLCSSQFRG 960  
 |||||||  
 Db 1391 FECEOKEENQGNESNIKRVQTVNITAGFPVVGQDKPVDNAKCSIKGSSRFLCSSQFRG 1450

```

QY 961 NETGLTPKHKGLLQNPYRIPLPLPIKSEYKTKCKNNLEENFEHSMSPEREMENIP 1020
DB 1451 NETGLTPKHKGLLQNPYRIPLPLPIKSEYKTKCKNNLEENFEHSMSPEREMENIP 1510
QY 1021 STVSTSRNNIRENVEKKAASSNNINEVGSSSTNEVGSSINIGSSDNIQALCRNRGPKL 1080
DB 1511 STVSTSRNNIRENVEKKAASSNNINEVGSSSTNEVGSSINIGSSDNIQALCRNRGPKL 1570
QY 1081 NAMLRGLVQLPEVYKOSLPGSNCKHPEIKKOEYEEVQVNTDPSFLLSDNLEOPMGSS 1140
DB 1571 NAMLRGLVQLPEVYKOSLPGSNCKHPEIKKOEYEEVQVNTDPSFLLSDNLEOPMGSS 1630
QY 1141 HASQVCSPTDLDLDDGETKEDTSFAENDIKESSAVFSKSVKQGLSRSPDPTHTHIAQ 1200
DB 1631 HASQVCSPTDLDLDDGETKEDTSFAENDIKESSAVFSKSVKQGLSRSPDPTHTHIAQ 1690
QY 1201 GYRRGAKKLESEENINSEDEELPCFQHLIFGKVNPIPSSQTPHSHVATRECSKNNEEN 1260
DB 1691 GYRRGAKKLESEENINSEDEELPCFQHLIFGKVNPIPSSQTPHSHVATRECSKNNEEN 1750
QY 1261 LSLKNSLNDSCNQVILAKASOEHLSEETKCSASLFSQCSLEEDLTANTQDPLIGS 1320
DB 1751 LSLKNSLNDSCNQVILAKASOEHLSEETKCSASLFSQCSLEEDLTANTQDPLIGS 1810
QY 1321 SKQMHQSTSGVGLSDKELVSDDERGTGLEBNNOEODSMDNLGEAASGCESETSVSE 1380
DB 1811 SKQMHQSTSGVGLSDKELVSDDERGTGLEBNNOEODSMDNLGEAASGCESETSVSE 1870
QY 1381 DCSGSSQSDILTTQORDTMOHNLTKLOEAMAELEVLFOHRSOPNSVPTIISDSSALE 1440
DB 1871 DCSGSSQSDILTTQORDTMOHNLTKLOEAMAELEVLFOHRSOPNSVPTIISDSSALE 1930
QY 1441 DLRNEQSTSEKAVLTQSKSESEYPIQONEPEGLSADKFEVSADSTSKNNEPEVERSSPSK 1500
DB 1931 DLRNEQSTSEKAVLTQSKSESEYPIQONEPEGLSADKFEVSADSTSKNNEPEVERSSPSK 1990
QY 1501 CPSLDDRWYMHSCSSGLQNRNTPSOBELIKVVDVEEQULEESGPHDLETSTLPRODLEG 1560
DB 1991 CPSLDDRWYMHSCSSGLQNRNTPSOBELIKVVDVEEQULEESGPHDLETSTLPRODLEG 2050
QY 1561 TPYLESGISLFSDDPESDPESEDRAPEASRVGNIPSTSLAKVPOLKVAESASPAAHNT 1620
DB 2051 TPYLESGISLFSDDPESDPESEDRAPEASRVGNIPSTSLAKVPOLKVAESASPAAHNT 2110
QY 1621 DTAGYNAMEESVSRKPELTASTERYNKRMSWVVSGLAPEEFMLVYKFKARKHHTLTNLI 1680
DB 2111 DTAGYNAMEESVSRKPELTASTERYNKRMSWVVSGLAPEEFMLVYKFKARKHHTLTNLI 2170
QY 1681 TEETHHYVAKTDAEFVCEERTLKYPFLGIAGKVVVSYFWVTOSIKERKMLNEHDFEYRGV 1740
DB 2171 TEETHHYVAKTDAEFVCEERTLKYPFLGIAGKVVVSYFWVTOSIKERKMLNEHDFEYRGV 2230
QY 1741 VNGRNHOGPKRARESDOKRIFRGLEICCYGFTNMPQOLEMVMVLCGASVYKELSSFTL 1800
DB 2231 VNGRNHOGPKRARESDOKRIFRGLEICCYGFTNMPQOLEMVMVLCGASVYKELSSFTL 2290
QY 1801 GTGVHPYVQPPDAMTENDNGFHAIGQCEAPVYTRRWLDSVALYQCELDYLLIPIQIPH 1860
DB 2291 GTGVHPYVQPPDAMTENDNGFHAIGQCEAPVYTRRWLDSVALYQCELDYLLIPIQIPH 2350
QY 1861 SHY 1863
DB 2351 SHY 2353

```

RESULT 9  
AAR81490 standard; Protein: 1863 AA.

ID AAR81490:  
AC AAR81490:  
XX  
DT 01-OCT-1996 (first entry)

```

XX DE BRCA1 mutant from patient 98.
XX KW Cancer therapy: breast and ovarian cancer predisposing gene; immunogen:
XX KM antibody production; germline alteration; probe; lesion neoplasia; human;
XX KW gene therapy; protein replacement therapy; protein mimetic; BRCA1.
XX OS Homo sapiens.
XX FH Key
XX FT Misc-difference 271
XX PD /note="V271M"
XX PN W09605306-A2.
XX PD 22-FEB-1996.
XX PF 11-AUG-1995: 95WO-US10202.
XX PR 07-JUN-1995: 95US-0483553.
XX PR 12-AUG-1994: 94US-0289221.
XX PR 02-SEP-1994: 94US-0300266.
XX PR 16-SEP-1994: 94US-0308104.
XX PR 29-NOV-1994: 94US-0348824.
XX PR 24-MAR-1995: 94US-0409305.
XX PR 07-JUN-1995: 95US-0480784.
XX PA (MYRI-) MYRIAD GENETICS INC.
XX PA (CANC-) CANCER INST.
XX PA (RECH-) CENT RECH DU CHUL.
XX PI Durocher F, Emi M, Nakamura Y, Simard J, Shattuck-Eidens DM;
XX DR WPI: 1996-139702/14.
XX DR N-PSDB: AAT17446.
XX PT New nucleic acid and polypeptide for mutant or polymorphic BRCA1
XX PS gene - for diagnosis and therapy of human breast and ovarian cancer
XX PS and for diagnosing pre-disposition to these cancers
XX PS Claim 1: 218bp; English.
XX CC AAR81483-R81497 and AAR81499-R81546 represent mutations of the protein
XX CC encoded by the human breast and ovarian cancer predisposing gene (BRCA1)
XX CC (see AAR81481 for wild type protein). These mutations can be used as
XX CC immunogens for antibody production. The mutant BRCA1 genes encoding
XX CC these sequences have at least 1 mutation or polymorphism in comparison
XX CC to the wild type cDNA (see AAT17438 for wild type). By detecting a
XX CC germline alteration in the wild type BRCA1 gene, a predisposition for
XX CC breast and ovarian cancer can be diagnosed. In one method, BRCA1 mRNA
XX CC isolated from a tissue sample from a subject has a probe, corresponding
XX CC to a fragment of the cDNA encoding the wild type BRCA1 sequence (or an
XX CC allele-specific probe for a mutation of it), added to it. The conditions
XX CC allow for hybridisation of the probe to the mRNA, and any hybridisation
XX CC which occurs is detected. Alternatively the BRCA1 gene in the tissue
XX CC sample is isolated, and a shift in electrophoretic mobility of single
XX CC stranded DNA from the sample on a non-denaturing polyacrylamide gel
XX CC indicates a mutation. These methods of detection can also diagnose a
XX CC lesion neoplasia associated with the BRCA1 locus. The methods may be
XX CC used in gene therapy, protein replacement therapy and protein mimetics,
XX CC and may be used to screen for drugs in cancer therapy.
XX SO Sequence 1863 AA:

```

Query Match 99.9%; Score 9639; DB 17; Length 1863;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1861; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 MDLSALRYEEVQYNNINAMOKITLPCICLLEIKPEVSTKCHIFCKPMLKLNOKKGPSO 60
DB 1 MDLSALRYEEVQYNNINAMOKITLPCICLLEIKPEVSTKCHIFCKPMLKLNOKKGPSO 60
QY 61 CPLCKNDITKRSIOESTRFSQVLEELIKITICAFQDGTGLEVANSYNFAKKNNSPEHLKD 120

```

```

|||||
Db 61 CPLCKNDITRSLOESTRFSQLVTEELLKIIICAFOLDITGLEVANSYNFAKKENSPHEHLKD 120
Qy 121 EVSTIIOSMGYRNRAKRLLOSEPNSIQETSLSVOLNLGTVRLTRKRIQOPKTSVYI 180
Db 121 EVSTIIOSMGYRNRAKRLLOSEPNSIQETSLSVOLNLGTVRLTRKRIQOPKTSVYI 180
Qy 181 ELGSDSESDYVNAKATYCSVGDQELLQITPOGTREIISLSAKKACESESDVNTNTEHQ 240
Db 181 ELGSDSESDYVNAKATYCSVGDQELLQITPOGTREIISLSAKKACESESDVNTNTEHQ 240
Qy 241 PSNNDLNTTEKRAAERHPEKYOSSVSNLHVEPCGTNTHASLOHENSLLITKDRNVE 300
Db 241 PSNNDLNTTEKRAAERHPEKYOSSVSNLHVEPCGTNTHASLOHENSLLITKDRNVE 300
Qy 301 KAEFCNKSQOPGLARSOHNWAGSKETCNDRRTSTETKAYDLNADPLCEKREMNKOKLPC 360
Db 301 KAEFCNKSQOPGLARSOHNWAGSKETCNDRRTSTETKAYDLNADPLCEKREMNKOKLPC 360
Qy 361 SENRDTEDYVWITLNGSIOKVNMFSRDELLSGSDSHGSESNKAVADVLVNEVD 420
Db 361 SENRDTEDYVWITLNGSIOKVNMFSRDELLSGSDSHGSESNKAVADVLVNEVD 420
Qy 421 EYSGSSSEKIDLLASDPHEALICKSERVHSKSVESNIEDKIFGKTYRRKASLPNLSHTYEN 480
Db 421 EYSGSSSEKIDLLASDPHEALICKSERVHSKSVESNIEDKIFGKTYRRKASLPNLSHTYEN 480
Qy 481 LIIAFYTEPOIIOERPLTNKLRKRRTSGLHPEDTICKADLAVOKTPEKINGTQOTE 540
Db 481 LIIAFYTEPOIIOERPLTNKLRKRRTSGLHPEDTICKADLAVOKTPEKINGTQOTE 540
Qy 541 ONGQVMNITNGSHENKTKGDSIOENEKNPNPIESLEKESAFKTKAEPISSTISNNELEINI 600
Db 541 ONGQVMNITNGSHENKTKGDSIOENEKNPNPIESLEKESAFKTKAEPISSTISNNELEINI 600
Qy 601 HNSKAPKNRLRRKSSTRHIIHLELVYSRNLSPNCTELOIDCSSSEELIKKKYKNOMPV 660
Db 601 HNSKAPKNRLRRKSSTRHIIHLELVYSRNLSPNCTELOIDCSSSEELIKKKYKNOMPV 660
Qy 661 RHSNIIOLMGEKEPATATKAKSKNKNNEOTSKRHDSDTPELKLITNAPSFTKCSMTSELKE 720
Db 661 RHSNIIOLMGEKEPATATKAKSKNKNNEOTSKRHDSDTPELKLITNAPSFTKCSMTSELKE 720
Qy 721 FVNSPLPREKEEKELETVKVSNNNAEDPKDMLSGERYLOTERSVESSTISLVPCTDGTQ 780
Db 721 FVNSPLPREKEEKELETVKVSNNNAEDPKDMLSGERYLOTERSVESSTISLVPCTDGTQ 780
Qy 781 ESISLLEVSTLGRKATEPNKCVSQAFAENPKGLIHGCSKDNRDTEGFKYPLIGHEVNH 840
Db 781 ESISLLEVSTLGRKATEPNKCVSQAFAENPKGLIHGCSKDNRDTEGFKYPLIGHEVNH 840
Qy 841 RETSIEMESELDAQIYONTFKYSKROSPALFSNPGNAEECAFTSAHSGLKQSPKVT 900
Db 841 RETSIEMESELDAQIYONTFKYSKROSPALFSNPGNAEECAFTSAHSGLKQSPKVT 900
Qy 901 FECQKEENOGKNBSNIKPVOTVITAGFPVYGOKDPVDNAKCSIGGSFECCLSSQPRG 960
Db 901 FECQKEENOGKNBSNIKPVOTVITAGFPVYGOKDPVDNAKCSIGGSFECCLSSQPRG 960
Qy 961 NETGLITPNKHGLLONDIYRIPLPIKSFYKTKCKNULNEEFHSMSPERENGNIPI 1020
Db 961 NETGLITPNKHGLLONDIYRIPLPIKSFYKTKCKNULNEEFHSMSPERENGNIPI 1020
Qy 1021 STVSTIRNNIRRENVFEASSNINEVGSSINELGSSDENIQAELGNRGPKL 1080
Db 1021 STVSTIRNNIRRENVFEASSNINEVGSSINELGSSDENIQAELGNRGPKL 1080
Qy 1081 NAMRLGLVLOPEVYKOLPGSNCKHPEIKKOEYEVOVOTVNTDPSPLISNLEOPQSS 1140
Db 1081 NAMRLGLVLOPEVYKOLPGSNCKHPEIKKOEYEVOVOTVNTDPSPLISNLEOPQSS 1140
Qy 1141 HASQVCEPDDLLDDEIKEIDTSFAENDIKESSAVFSKVOGELSRSPSPFTHTHLAQ 1200
Db 1141 HASQVCEPDDLLDDEIKEIDTSFAENDIKESSAVFSKVOGELSRSPSPFTHTHLAQ 1200

```

```

Db 1141 HASQVCEPDDLLDDEIKEIDTSFAENDIKESSAVFSKVOGELSRSPSPFTHTHLAQ 1200
Qy 1201 GYRGAKKLESSSEENLSSDEBELPCFQHLIFGKYNNTIPSGSTRSTYATBELSKNTENL 1260
Db 1201 GYRGAKKLESSSEENLSSDEBELPCFQHLIFGKYNNTIPSGSTRSTYATBELSKNTENL 1260
Qy 1261 LSLKNSLNDCSNOYIILAKASOEHLSEETKSASALFSQCELEDLTANTTOOPFLIGS 1320
Db 1261 LSLKNSLNDCSNOYIILAKASOEHLSEETKSASALFSQCELEDLTANTTOOPFLIGS 1320
Qy 1321 SKQMRHOSQSVGLSKELVSDDERGTGLENNQEOBSMDNIGEAASCESETSVSE 1380
Db 1321 SKQMRHOSQSVGLSKELVSDDERGTGLENNQEOBSMDNIGEAASCESETSVSE 1380
Qy 1381 DCSGLSSQSDILTTQOQRTMOHNLIKIQOEAELEAVLEQHSOPSNYSPIISDSSALE 1440
Db 1381 DCSGLSSQSDILTTQOQRTMOHNLIKIQOEAELEAVLEQHSOPSNYSPIISDSSALE 1440
Qy 1441 DLRNPEOSTSEKAVLTQSSEYPISONPEGLISADKFEVSADSSSTSKNKEGVRSRSPSK 1500
Db 1441 DLRNPEOSTSEKAVLTQSSEYPISONPEGLISADKFEVSADSSSTSKNKEGVRSRSPSK 1500
Qy 1501 CPSLDDRWYMHSCGSLQNRNYPQOEELIKVYVEEQOLEESGPHDLTETSYLPRODLEG 1560
Db 1501 CPSLDDRWYMHSCGSLQNRNYPQOEELIKVYVEEQOLEESGPHDLTETSYLPRODLEG 1560
Qy 1561 TPYLESGISLFSDDPESDPESDRAPESARVGNIPSSITALKVPOLKVAESQSPAAHATT 1620
Db 1561 TPYLESGISLFSDDPESDPESDRAPESARVGNIPSSITALKVPOLKVAESQSPAAHATT 1620
Qy 1621 DTAGYNAMBEESVSEKPELTASTERYNKRMSMNVYSGLTPEEFMLVYFARKHHTITNLII 1680
Db 1621 DTAGYNAMBEESVSEKPELTASTERYNKRMSMNVYSGLTPEEFMLVYFARKHHTITNLII 1680
Qy 1681 TEETTHVVMKTDAEVCERTLKYELGIAGCKWVSYTFWVTOISIKERMLNHEDEVRGDV 1740
Db 1681 TEETTHVVMKTDAEVCERTLKYELGIAGCKWVSYTFWVTOISIKERMLNHEDEVRGDV 1740
Qy 1741 VNGRNHOGPKRARBSOBKRTFRGLCTCYGFTWMPDOLLEMYVOLCGASVYKELSSFTL 1800
Db 1741 VNGRNHOGPKRARBSOBKRTFRGLCTCYGFTWMPDOLLEMYVOLCGASVYKELSSFTL 1800
Qy 1801 GTGVHPIVVOVPODAMTEDENGFHAIGOMCEAPVTRREWLDSVALYOCOEIDTYLIPQIPH 1860
Db 1801 GTGVHPIVVOVPODAMTEDENGFHAIGOMCEAPVTRREWLDSVALYOCOEIDTYLIPQIPH 1860
Qy 1861 SHY 1863
Db 1861 SHY 1863

RESULT 10
AAR81500
ID AAR81500 standard; Protein; 1863 AA.
XX
AC AAR81500;
XX
DT 02-OCT-1996 (first entry)
XX
DE BRCA1 mutant from sample sets YN98 and YN7.
XX
KW Cancer therapy; breast and ovarian cancer predisposing gene; immunogen;
KW antibody production; germline alteration; probe; lesion neoplasia; human;
KW gene therapy; protein replacement therapy; protein mimetic; BRCA1.
XX
OS Homo sapiens.
XX
FH Key location/Qualifiers
FT Misc-difference 271
XX /note- "v271m"
XX
PN W09605306-A2.
XX

```





Db 1381 DCGSLSSQSDILTTQOQDMQHNLIKQOEMAELEAVLEQHGQSPNSYPSIISSDALE 1440  
 QY 1441 DLNRPQOSTSEKAVLSQKSSSEPIQONPEGLSADFEVADSTSTKNKPGVERSSPSK 1500  
 Db 1441 DLNRPQOSTSEKAVLSQKSSSEPIQONPEGLSADFEVADSTSTKNKPGVERSSPSK 1500  
 QY 1501 CPSLIDRWYHSCSGSLQNNRNPQOELLKVVDVEEQOLEESGPHDLTETSYLPRODLEG 1560  
 Db 1501 CPSLIDRWYHSCSGSLQNNRNPQOELLKVVDVEEQOLEESGPHDLTETSYLPRODLEG 1560  
 QY 1561 TPLLESGISLFSDDPESDPEADAPASARVGNIPSSTSALKVQOLKYAESAGSPAAAHHT 1620  
 Db 1561 TPLLESGISLFSDDPESDPEADAPASARVGNIPSSTSALKVQOLKYAESAGSPAAAHHT 1620  
 QY 1621 DTGYNAMSESVSRPELTASTERYNKRMSVSGLTPEEFMLVYKFAKHHTLTNL 1680  
 Db 1621 DTGYNAMSESVSRPELTASTERYNKRMSVSGLTPEEFMLVYKFAKHHTLTNL 1680  
 QY 1681 TEETTHVYMKTDAEFYCEERTLKLYELGIAGKMWVSFYWYQSIKERKMLNEHDFEVRGDY 1740  
 Db 1681 TEETTHVYMKTDAEFYCEERTLKLYELGIAGKMWVSFYWYQSIKERKMLNEHDFEVRGDY 1740  
 QY 1741 VGNRNHOGPKRARESDRKTFRGLEICCYGPTNMPDLEMMVQLGASVKEISSFTL 1800  
 Db 1741 VGNRNHOGPKRARESDRKTFRGLEICCYGPTNMPDLEMMVQLGASVKEISSFTL 1800  
 QY 1801 GTGVHPIVYQPDAMTDENGFHAIGOMCEAPVYTRREMVLDVALYCCQELDTYLLIPIQIRH 1860  
 Db 1801 GTGVHPIVYQPDAMTDENGFHAIGOMCEAPVYTRREMVLDVALYCCQELDTYLLIPIQIRH 1860  
 QY 1861 SHY 1863  
 Db 1861 SHY 1863

RESULT 11  
 AAR81522 standard; Protein; 1863 AA.  
 AAR81522:  
 02-OCT-1996 (first entry)  
 BRCA1 mutant from sample set MSK12871.  
 Cancer therapy: breast and ovarian cancer predisposing gene; immunogen;  
 antibody production; germline alteration; probe; lesion neoplasia; human;  
 gene therapy; protein replacement therapy; protein mimetic; BRCA1.  
 Homo sapiens.  
 Key Location/Qualifiers  
 Misc-difference 1219 /note= "E1219D"  
 W09605306-AZ.  
 22-FEB-1996.  
 11-AUG-1995; 95MO-US10202.  
 07-JUN-1995; 95US-0483553.  
 12-AUG-1994; 94US-0289221.  
 16-SEP-1994; 94US-0300266.  
 29-NOV-1994; 94US-0308104.  
 24-MAR-1995; 94US-0348824.  
 07-JUN-1995; 95US-0409305.  
 (MYRI-) MYRIAD GENETICS INC.  
 (CANC-) CANCER INST.  
 (RECH-) CENT RECH DU CHUL.

XX Durocher F, Emi M, Nakamura Y, Simard J, Shattuck-Bidens DM;  
 PI WPI: 1996-139702/14.  
 DR N-PSDB; AAT17479.  
 XX New nucleic acid and polypeptide for mutant or polymorphic BRCA1  
 PT gene - for diagnosis and therapy of human breast and ovarian cancer  
 PT and for diagnosing pre-disposition to these cancers  
 XX  
 PS Claim 1; : 218pp: English.  
 XX  
 CC AAR81483-R81497 and AAR81499-R81546 represent mutations of the protein  
 CC encoded by the human breast and ovarian cancer predisposing gene (BRCA1)  
 CC (see AAR81481 for wild type protein). These mutations can be used as  
 CC immunogens for antibody production. The mutant BRCA1 genes encoding  
 CC these sequences have at least 1 mutation or polymorphism in comparison  
 CC to the wild type cDNA (see AAT17438 for wild type). By detecting a  
 CC germline alteration in the wild type BRCA1 gene, a predisposition for  
 CC breast and ovarian cancer can be diagnosed. In one method, BRCA1 mRNA  
 CC isolated from a tissue sample from a subject has a probe, corresponding  
 CC to a fragment of the cDNA encoding the wild type BRCA1 sequence (or an  
 CC allele-specific probe for a mutation of it), added to it. The conditions  
 CC allow for hybridisation of the probe to the mRNA, and any hybridisation  
 CC which occurs is detected. Alternatively the BRCA1 gene in the tissue  
 CC sample is isolated, and a shift in electrophoretic mobility of single  
 CC stranded DNA from the sample on a non-denaturing polyacrylamide gel  
 CC indicates a mutation. These methods of detection can also diagnose a  
 CC lesion neoplasia associated with the BRCA1 locus. The methods may be  
 CC used in gene therapy, protein replacement therapy and protein mimetics,  
 CC and may be used to screen for drugs in cancer therapy.  
 CC  
 SQ Sequence 1863 AA:  
 Query Match 99.9%; Score 9639; DB 17; Length 1863;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1861; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MDLSALRVEVQVNIAMOKILECPICLEIKRPVSTKCHIFCKRCKMLLNOKKGPSP 60  
 Db 1 MDLSALRVEVQVNIAMOKILECPICLEIKRPVSTKCHIFCKRCKMLLNOKKGPSP 60  
 QY 61 CPLCKNDITKRSLOESTRESQVLEELIKITICAFQDLTGLEYANSNYFAKKENNSPEHLND 120  
 Db 61 CPLCKNDITKRSLOESTRESQVLEELIKITICAFQDLTGLEYANSNYFAKKENNSPEHLND 120  
 QY 121 EVSIIOSMGYRNRAKRLLOSEPENPSLOETSLSVOLSNIGVYRTLKRIOPQKTSVYI 180  
 Db 121 EVSIIOSMGYRNRAKRLLOSEPENPSLOETSLSVOLSNIGVYRTLKRIOPQKTSVYI 180  
 QY 181 ELGSDSSEPTVNAKATYCSVGDOELLOITPGTFDELSLSAKKAACEFSETDTNTEHHO 240  
 Db 181 ELGSDSSEPTVNAKATYCSVGDOELLOITPGTFDELSLSAKKAACEFSETDTNTEHHO 240  
 QY 241 PSNNDLNTTEKRAERHPRKYOSSYSNLHVEPCGNTNTHASSLOHENSLLTKDRMANE 300  
 Db 241 PSNNDLNTTEKRAERHPRKYOSSYSNLHVEPCGNTNTHASSLOHENSLLTKDRMANE 300  
 QY 301 KAEFCNKSQOPGLARQHNRMWAGSKETCDNRTPSTTEKRVYDLNADPLCERKEKNNOKLPC 360  
 Db 301 KAEFCNKSQOPGLARQHNRMWAGSKETCDNRTPSTTEKRVYDLNADPLCERKEKNNOKLPC 360  
 QY 361 SENPRDTEVPMTLLSSIOKVNEMFSRSDDELGSDSDSHGSESNAAKVAADVLDVINEVD 420  
 Db 361 SENPRDTEVPMTLLSSIOKVNEMFSRSDDELGSDSDSHGSESNAAKVAADVLDVINEVD 420  
 QY 421 EYSGSSEKIDLLASDPHEALICKSERVHSHSVESNIEDKIFPKTYRKKASLNLNLSHVEN 480  
 Db 421 EYSGSSEKIDLLASDPHEALICKSERVHSHSVESNIEDKIFPKTYRKKASLNLNLSHVEN 480  
 QY 481 LITGAFVTEPQIIOERPLTNLKKRRRPTSGLHPEDFIKADLAVOKTPEMINOCTNOTE 540  
 Db 481 LITGAFVTEPQIIOERPLTNLKKRRRPTSGLHPEDFIKADLAVOKTPEMINOCTNOTE 540

QY 541 QNGOVNITNSGHENKTGDSIONENKPNPIESLEKESAFKTKAEPDISSISIMMELEANI 600  
 |||||  
 Db 541 QNGOVNITNSGHENKTGDSIONENKPNPIESLEKESAFKTKAEPDISSISIMMELEANI 600  
 QY 601 HNSKAPKRRRLRRKSSRTIHALELVYRNLSPPNCTELOIDSCSSSEETKRRKYOMPV 660  
 |||||  
 Db 601 HNSKAPKRRRLRRKSSRTIHALELVYRNLSPPNCTELOIDSCSSSEETKRRKYOMPV 660  
 QY 661 RRSRNLOLECKEPATGAKSKNPNQOTSRHSDPFPELKLINAGSGFTKCSNTSELKE 720  
 |||||  
 Db 661 RRSRNLOLECKEPATGAKSKNPNQOTSRHSDPFPELKLINAGSGFTKCSNTSELKE 720  
 QY 721 FVNPSLPREEKEKLETYKVSNNADPKDMLSGERYLOTERSVSSSISLVGTDTYGT 780  
 |||||  
 Db 721 FVNPSLPREEKEKLETYKVSNNADPKDMLSGERYLOTERSVSSSISLVGTDTYGT 780  
 QY 781 ESISLLEVSTLGAKTEPKKCVSQAFAFENPKGLHGCSKDNNDTEGFKYPLGHEVNH 840  
 |||||  
 Db 781 ESISLLEVSTLGAKTEPKKCVSQAFAFENPKGLHGCSKDNNDTEGFKYPLGHEVNH 840  
 QY 841 RETSIEMESELDAOYLONTFKYKROSFALFSPGNAEPECATFSAHSGSLKKOSPKYT 900  
 |||||  
 Db 841 RETSIEMESELDAOYLONTFKYKROSFALFSPGNAEPECATFSAHSGSLKKOSPKYT 900  
 QY 901 FECEOKREENOGKNESNIKPVQTVNITAGFPVVGQKDPVDNAKCSIKGSRFCLSSQFRG 960  
 |||||  
 Db 901 FECEOKREENOGKNESNIKPVQTVNITAGFPVVGQKDPVDNAKCSIKGSRFCLSSQFRG 960  
 QY 961 NETGLITPKKHGLONPRIPPLPIKSYKTKCKNLLFENFEHMSGPERMGNDTP 1020  
 |||||  
 Db 961 NETGLITPKKHGLONPRIPPLPIKSYKTKCKNLLFENFEHMSGPERMGNDTP 1020  
 QY 1021 STVSTISRNINIRENFKEASSSININEVGSSSTNEVGSSINIEGSSDNIOALGRNNGPKL 1080  
 |||||  
 Db 1021 STVSTISRNINIRENFKEASSSININEVGSSSTNEVGSSINIEGSSDNIOALGRNNGPKL 1080  
 QY 1081 NAMRLGLVLOPEYVKOSLPGSNCKHPEIKKOEYEEVQTVNDFSPLYLSIDNLEQPMGSS 1140  
 |||||  
 Db 1081 NAMRLGLVLOPEYVKOSLPGSNCKHPEIKKOEYEEVQTVNDFSPLYLSIDNLEQPMGSS 1140  
 QY 1141 HASOVCEETPDLLDGEIKEDTSFAENDIKESSANFESKSVQCGELSRSPPTHTHIAQ 1200  
 |||||  
 Db 1141 HASOVCEETPDLLDGEIKEDTSFAENDIKESSANFESKSVQCGELSRSPPTHTHIAQ 1200  
 QY 1201 GYRRGAKKLESSEENLSEDEELPCFOHLFGKVNNIPOSOTRHSVATECLSKNTEENI 1260  
 |||||  
 Db 1201 GYRRGAKKLESSEENLSEDEELPCFOHLFGKVNNIPOSOTRHSVATECLSKNTEENI 1260  
 QY 1261 LSLKNSLNDCSNOVILAKASQEHHLSEETKCSASLFSQCSLEEDLTJANTNDPPLIGS 1320  
 |||||  
 Db 1261 LSLKNSLNDCSNOVILAKASQEHHLSEETKCSASLFSQCSLEEDLTJANTNDPPLIGS 1320  
 QY 1321 SKQMRHOSQGVGLSDKELVSDDERGCTGLEENNOEDOSMSNLCEAASGCSETSVSE 1380  
 |||||  
 Db 1321 SKQMRHOSQGVGLSDKELVSDDERGCTGLEENNOEDOSMSNLCEAASGCSETSVSE 1380  
 QY 1381 DCSGSLSSOSDILITTOORDTMOHNLKLOQEMAELEVLBOHGSOPNSYPSIISDSALE 1440  
 |||||  
 Db 1381 DCSGSLSSOSDILITTOORDTMOHNLKLOQEMAELEVLBOHGSOPNSYPSIISDSALE 1440  
 QY 1441 DLNRNEQOSTSEKAVILTSOKSEYPISQNPEGLSADKFEVSADSTSKNKEPVERSPSK 1500  
 |||||  
 Db 1441 DLNRNEQOSTSEKAVILTSOKSEYPISQNPEGLSADKFEVSADSTSKNKEPVERSPSK 1500  
 QY 1501 CPSLDDRRMYHSCSSGLONRNPPOEELIKVYDVEEQOLEESGPHDLTETSTYLPDOLDG 1560  
 |||||  
 Db 1501 CPSLDDRRMYHSCSSGLONRNPPOEELIKVYDVEEQOLEESGPHDLTETSTYLPDOLDG 1560  
 QY 1561 TPYLESGISLFGDDESPDESEDRAPESARVGNIPSSTSALKVPOLKVAESAGSPAAAHHT 1620  
 |||||  
 Db 1561 TPYLESGISLFGDDESPDESEDRAPESARVGNIPSSTSALKVPOLKVAESAGSPAAAHHT 1620

QY 1621 DTAGYNAMEESVREKPELTASTERYNKRMSVSGLTPEEFMLYKFAFRHHITLNL 1680  
 |||||  
 Db 1621 DTAGYNAMEESVREKPELTASTERYNKRMSVSGLTPEEFMLYKFAFRHHITLNL 1680  
 QY 1681 TEETTHVAKTDAEFCERTLKYFGLIAGKRVVSYFWTOSIKRRKMLNHEDEVRGDV 1740  
 |||||  
 Db 1681 TEETTHVAKTDAEFCERTLKYFGLIAGKRVVSYFWTOSIKRRKMLNHEDEVRGDV 1740  
 QY 1741 VNGRNHOGPKRRAESODRKIFRGLEICCYGPTNMPDOLBMMVOLCGASVYKELSPFL 1800  
 |||||  
 Db 1741 VNGRNHOGPKRRAESODRKIFRGLEICCYGPTNMPDOLBMMVOLCGASVYKELSPFL 1800  
 QY 1801 GTGVHPITVVQPDWATEDNGFHAIGOMCEAPVYTRBWLDSVALYQCOELDTYLIPI 1860  
 |||||  
 Db 1801 GTGVHPITVVQPDWATEDNGFHAIGOMCEAPVYTRBWLDSVALYQCOELDTYLIPI 1860  
 QY 1861 SHY 1863  
 |||  
 Db 1861 SHY 1863  
 |||

RESULT 12  
 AAR81536  
 ID AAR81536 standard; Protein; 1863 AA.  
 AC AAR81536;  
 DT 02-OCT-1996 (first entry)  
 XX  
 DE BRCA1 mutant from PM07.  
 KW Cancer therapy; breast and ovarian cancer predisposing gene; immunogen;  
 KW antibody production; germline alteration; probe; lesion neoplasia; human;  
 KW gene therapy; protein replacement therapy; protein mimetic; BRCA1.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 1183  
 FT /note= "K1183"  
 XX  
 PN W09605306-A2.  
 XX  
 PD 22-FEB-1996.  
 XX  
 PF 11-AUG-1995; 95MO-US10202.  
 XX  
 PR 07-JUN-1995; 95US-0483553.  
 PR 12-AUG-1994; 94US-0289221.  
 PR 02-SEP-1994; 94US-0300266.  
 PR 16-SEP-1994; 94US-0308104.  
 PR 29-NOV-1994; 94US-0348824.  
 PR 24-MAR-1995; 95US-0409305.  
 PR 07-JUN-1995; 95US-0480784.  
 XX  
 PA (MYRI-) MYRIAD GENETICS INC.  
 PA (CANC-) CANCER INST.  
 PA (RECH-) CENT RECH DU CHUL.  
 XX  
 PI Durocher F, Emi M, Nakamura Y, Simard J, Shattuck-Eidens DM;  
 XX  
 DR WPI: 1996-139702/14.  
 DR N-PSDB: AAT17494.  
 XX  
 PT New nucleic acid and polypeptide for mutant or polymorphic BRCA1  
 PT gene - for diagnosis and therapy of human breast and ovarian cancer  
 PT and for diagnosing pre-disposition to these cancers  
 XX  
 PS Claim 1; : 218pp; English.  
 CC AAR81483-R81497 and AAR81499-R81546 represent mutations of the protein  
 CC encoded by the human breast and ovarian cancer predisposing gene (BRCA1)  
 CC (see AAR81481 for wild type protein). These mutations can be used as

CC immunogens for antibody production. The mutant BRCA1 genes encoding  
CC these sequences have at least 1 mutation or polymorphism in comparison  
CC to the wild type cDNA (see AAT17438 for wild type). By detecting a  
CC germline alteration in the wild type BRCA1 gene, a predisposition for  
CC breast and ovarian cancer can be diagnosed. In one method, BRCA1 mRNA  
CC isolated from a tissue sample from a subject has a probe, corresponding  
CC to a fragment of the cDNA encoding the wild type BRCA1 sequence (or an  
CC allele-specific probe for a mutation of it), added to it. The conditions  
CC allow for hybridisation of the probe to the mRNA, and any hybridisation  
CC which occurs is detected. Alternatively the BRCA1 gene in the tissue  
CC sample is isolated, and a shift in electrophoretic mobility of single  
CC stranded DNA from the sample on a non-denaturing polyacrylamide gel  
CC indicates a mutation. These methods of detection can also diagnose a  
CC lesion neoplasia associated with the BRCA1 locus. The methods may be  
CC used in gene therapy, protein replacement therapy and protein mimetics,  
CC and may be used to screen for drugs in cancer therapy.

XX Sequence 1863 AA:  
SQ

Query Match 99.9%; Score 9639; DB 17; Length 1863;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1861; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDLSALRVEEVQVNIAMOKILECPICLEIKPEVSTKCDHIFCKFCMKLLNQKKGPQ 60  
DB 1 MDLSALRVEEVQVNIAMOKILECPICLEIKPEVSTKCDHIFCKFCMKLLNQKKGPQ 60  
QY 61 CPLCKNDITKRSIQESTRSQVEELIKITICAFQDGLGEYANSYNPAKKENSPHKLND 120  
DB 61 CPLCKNDITKRSIQESTRSQVEELIKITICAFQDGLGEYANSYNPAKKENSPHKLND 120  
QY 121 EVSIIOSMGYRNRAKLLQSEPNPSIQESTLSVOLSNLGTVRTLTQKRIQPKTSVYI 180  
DB 121 EVSIIOSMGYRNRAKLLQSEPNPSIQESTLSVOLSNLGTVRTLTQKRIQPKTSVYI 180  
QY 121 EVSIIOSMGYRNRAKLLQSEPNPSIQESTLSVOLSNLGTVRTLTQKRIQPKTSVYI 180  
DB 121 EVSIIOSMGYRNRAKLLQSEPNPSIQESTLSVOLSNLGTVRTLTQKRIQPKTSVYI 180  
QY 181 ELGSDSSEDTVNAKTCYSGVDELQITPOGTRDEISLDSAKKAEFESETDVTNEHQ 240  
DB 181 ELGSDSSEDTVNAKTCYSGVDELQITPOGTRDEISLDSAKKAEFESETDVTNEHQ 240  
QY 241 PSNNNDITTEKRAEHRPEKYOGSSVSNLHVEPCGNTNTHASSIQHNSSLTLTKDMANVE 300  
DB 241 PSNNNDITTEKRAEHRPEKYOGSSVSNLHVEPCGNTNTHASSIQHNSSLTLTKDMANVE 300  
QY 301 KAEFCNKSKOPGLARSOHNMAKSKETCNDRTPTSTEEKYVDLADPLCEKKEKNNKQLPC 360  
DB 301 KAEFCNKSKOPGLARSOHNMAKSKETCNDRTPTSTEEKYVDLADPLCEKKEKNNKQLPC 360  
QY 301 KAEFCNKSKOPGLARSOHNMAKSKETCNDRTPTSTEEKYVDLADPLCEKKEKNNKQLPC 360  
DB 301 KAEFCNKSKOPGLARSOHNMAKSKETCNDRTPTSTEEKYVDLADPLCEKKEKNNKQLPC 360  
QY 361 SENPRTEDVPWITLNSSTOKVNEWFSRSDDELIGSDSHDGESENAKVAADVLDVINEVD 420  
DB 361 SENPRTEDVPWITLNSSTOKVNEWFSRSDDELIGSDSHDGESENAKVAADVLDVINEVD 420  
QY 421 EYSGSEKIDLLASDHEALICSEVHKSVEVSNIEDKIFGKTYRKKAISPLVSHVTEN 480  
DB 421 EYSGSEKIDLLASDHEALICSEVHKSVEVSNIEDKIFGKTYRKKAISPLVSHVTEN 480  
QY 481 LIIGAFVTEPQIIQERPLTNLTKRRRPTSGLHPEDFIKKADLAVOKTEPMINQGTNOTE 540  
DB 481 LIIGAFVTEPQIIQERPLTNLTKRRRPTSGLHPEDFIKKADLAVOKTEPMINQGTNOTE 540  
QY 541 QNGQVNNITNSGHEKTKGDSIONENPNPISLEKESAFKTAEDISSISIMLEELNI 600  
DB 541 QNGQVNNITNSGHEKTKGDSIONENPNPISLEKESAFKTAEDISSISIMLEELNI 600  
QY 601 HNSKAKKRLRKRKSTRIHHALELVYVSNLSPNCTELOIDSCSSSEETKKKKYOMYV 660  
DB 601 HNSKAKKRLRKRKSTRIHHALELVYVSNLSPNCTELOIDSCSSSEETKKKKYOMYV 660  
QY 661 RHSRNQIQLEMGKRPATGAKSKNPNQTSKRHSDPFPBLKLTNAGSFTKCSNTSELKE 720  
DB 661 RHSRNQIQLEMGKRPATGAKSKNPNQTSKRHSDPFPBLKLTNAGSFTKCSNTSELKE 720  
QY 721 FVNPSPLPREEKEKLETYKVYNNADPDKDMLSGERYLOTENSVESSSISLVPGTDYGTQ 780  
DB 721 FVNPSPLPREEKEKLETYKVYNNADPDKDMLSGERYLOTENSVESSSISLVPGTDYGTQ 780

DB 721 FVNPSPLPREEKEKLETYKVYNNADPDKDMLSGERYLOTENSVESSSISLVPGTDYGTQ 780  
QY 781 ESTSILEVSTLGRKAKPEPNKCVSOCAAFEMPKLJHGCSMDNNDPEGEFYPLGHEVNH 840  
DB 781 ESTSILEVSTLGRKAKPEPNKCVSOCAAFEMPKLJHGCSMDNNDPEGEFYPLGHEVNH 840  
QY 841 RETSIEMEESIELDQVLONFKFKSKROSFALFNPNGNAEECATFSAHSGSLKKQSPKV 900  
DB 841 RETSIEMEESIELDQVLONFKFKSKROSFALFNPNGNAEECATFSAHSGSLKKQSPKV 900  
QY 901 FECEQKEENQKKNESNIKPVQVNTAGFPVYQOKRPVDNAKCSIKGSRFCLSSQFRG 960  
DB 901 FECEQKEENQKKNESNIKPVQVNTAGFPVYQOKRPVDNAKCSIKGSRFCLSSQFRG 960  
QY 961 NEGLITPNKHGLQNPYRIRPLPIKSPKTKCKNLLLENEEHSMSPEREMGENTP 1020  
DB 961 NEGLITPNKHGLQNPYRIRPLPIKSPKTKCKNLLLENEEHSMSPEREMGENTP 1020  
QY 1021 STVSTISRNNIRBNVKEASSSNIINEVGSSTNEVGSSINEIGSSDENIQALGRNRPKL 1080  
DB 1021 STVSTISRNNIRBNVKEASSSNIINEVGSSTNEVGSSINEIGSSDENIQALGRNRPKL 1080  
QY 1081 NAMLRGLVLOPEVYKOSLPESNCKHPEIKQOEVEVQVNTDFSPYLLISDNEQPMGSS 1140  
DB 1081 NAMLRGLVLOPEVYKOSLPESNCKHPEIKQOEVEVQVNTDFSPYLLISDNEQPMGSS 1140  
QY 1141 HASOVSEPPDDLDDGELKEPDSFPAENDIKESAAFSKVSQGELSRSPPTHTHQAQ 1200  
DB 1141 HASOVSEPPDDLDDGELKEPDSFPAENDIKESAAFSKVSQGELSRSPPTHTHQAQ 1200  
QY 1201 GYRGAKKLTLESSSEENLSEDEDELPCFQHLIFGVNNIIPQSOTRHSVATCELSKNTNEEN 1260  
DB 1201 GYRGAKKLTLESSSEENLSEDEDELPCFQHLIFGVNNIIPQSOTRHSVATCELSKNTNEEN 1260  
QY 1261 LSLKNSLNDCSNOVILAKASQEHNLSEETKCSALFSSQCSLEDELTAANTQDPLIGS 1320  
DB 1261 LSLKNSLNDCSNOVILAKASQEHNLSEETKCSALFSSQCSLEDELTAANTQDPLIGS 1320  
QY 1321 SKOMRHOSEOGVGLSDKELVSDDEERGCTLENNQEDQSMONLCEAASGCCSEFSVSP 1380  
DB 1321 SKOMRHOSEOGVGLSDKELVSDDEERGCTLENNQEDQSMONLCEAASGCCSEFSVSP 1380  
QY 1381 DCSGLSSQSDILITQOQDRTMQHNLKLOQEMALEVLEQHGSPQNSVPSIISDSALE 1440  
DB 1381 DCSGLSSQSDILITQOQDRTMQHNLKLOQEMALEVLEQHGSPQNSVPSIISDSALE 1440  
QY 1441 DLNRPQOSTSEKAVLTQKSSSEYPISONPEGLSADKFEVSAOSTSKNKEPGEVRSPPSK 1500  
DB 1441 DLNRPQOSTSEKAVLTQKSSSEYPISONPEGLSADKFEVSAOSTSKNKEPGEVRSPPSK 1500  
QY 1501 CPSLDRRWYHSGSGSLQNRNYPISOEBELIKVYVVEEQOLEESGPHULTETSTYLP 1560  
DB 1501 CPSLDRRWYHSGSGSLQNRNYPISOEBELIKVYVVEEQOLEESGPHULTETSTYLP 1560  
QY 1561 TPYLESGISLFSDDPSPDSSEDRAPESARVGINIPSSSTSLKVPOLVAESAQSPA 1620  
DB 1561 TPYLESGISLFSDDPSPDSSEDRAPESARVGINIPSSSTSLKVPOLVAESAQSPA 1620  
QY 1621 DTAAGYNAMEESVREKPELTASTERYNKRMSVVSGLTPEEFMLVYKFAKHHITLTNLI 1680  
DB 1621 DTAAGYNAMEESVREKPELTASTERYNKRMSVVSGLTPEEFMLVYKFAKHHITLTNLI 1680  
QY 1681 TEFTTHVVKTAEFCEPRTKFLGIGAGKWWVSFWWTQSTKKEKMLNEHDFEVRGV 1740  
DB 1681 TEFTTHVVKTAEFCEPRTKFLGIGAGKWWVSFWWTQSTKKEKMLNEHDFEVRGV 1740  
QY 1741 VNGRNHOGKRAESODRKIFRGLETICGYPFTNMPDLOLEMMVQJCGASVYKELSSFTL 1800  
DB 1741 VNGRNHOGKRAESODRKIFRGLETICGYPFTNMPDLOLEMMVQJCGASVYKELSSFTL 1800  
QY 1801 GTGVNPIVVOVDPAWTEGNGFHAIGOMCAPVYTRBWLDSVALYOCQELDTYLLIPOIPH 1860  
DB 1801 GTGVNPIVVOVDPAWTEGNGFHAIGOMCAPVYTRBWLDSVALYOCQELDTYLLIPOIPH 1860

QY 1861 SHY 1863  
 |||  
 Db 1861 SHY 1863

RESULT 13  
 AAR81540  
 AAR81540 standard; Protein; 1863 AA.

XX AAR81540;  
 AC AAR81540;  
 XX 02-OCT-1996 (first entry)  
 DT  
 XX  
 DE BRCA1 mutant from PM20.  
 XX  
 KW Cancer therapy; breast and ovarian cancer predisposing gene; immunogen;  
 KW antibody production; germline alteration; probe; lesion neoplasia; human;  
 KW gene therapy; protein replacement therapy; protein mimetic; BRCA1.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 1040 /note="S1040N"  
 FT  
 XX  
 PN M09605306-A2.  
 XX  
 PD 22-FEB-1996.  
 XX  
 PF 11-AUG-1995; 95WO-US10202.  
 XX  
 PR 07-JUN-1995; 95US-0483553.  
 PR 12-AUG-1994; 94US-0289221.  
 PR 02-SEP-1994; 94US-0300266.  
 PR 16-SEP-1994; 94US-0308104.  
 PR 29-NOV-1994; 94US-0348824.  
 PR 24-MAR-1995; 95US-0409305.  
 PR 07-JUN-1995; 95US-0480784.  
 XX  
 PA (MYRI-) MYRIAD GENETICS INC.  
 PA (CANC-) CANCER INST.  
 PA (RECH-) CENT RECH DU CHUL.  
 XX  
 PI Durocher F, Eml M, Nakamura Y, Simard J, Shattuck-Eidens DM;  
 DR MPI: 1996-139702/14.  
 DR N-PSDB: AAT17502.  
 XX  
 PT New nucleic acid and polypeptide for mutant or polymorphic BRCA1  
 PT gene - for diagnosis and therapy of human breast and ovarian cancer  
 PT and for diagnosing pre-disposition to these cancers  
 XX  
 PS Claim 1; : 218pp: English.

CC and may be used to screen for drugs in cancer therapy.  
 XX  
 SQ Sequence 1863 AA;  
 QY Query Match 99.9%; Score 9639; DB 17; Length 1863;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1861; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDLSALRYEEVONVIMAKOILIECPICELIKEPYSTGDIHFCKFCLLNOKKPSQ 60  
 |||||||  
 Db 1 MDLSALRYEEVONVIMAKOILIECPICELIKEPYSTGDIHFCKFCLLNOKKPSQ 60

QY 61 CPUCNDITKRSLOESTREFSOLVBEELIKTICAFOLDGLLEVANSYNPAKKENSPHUKD 120  
 |||||||  
 Db 61 CPUCNDITKRSLOESTREFSOLVBEELIKTICAFOLDGLLEVANSYNPAKKENSPHUKD 120

QY 121 EVSIIOSMGYRNKARLLQSEPEPNSLOETSLVSQSLGTVRLTKRORIOPKTSVYI 180  
 |||||||  
 Db 121 EVSIIOSMGYRNKARLLQSEPEPNSLOETSLVSQSLGTVRLTKRORIOPKTSVYI 180

QY 181 ELGSDSSEDTYKATYCGVDOELQITPOGTRDEISLSAKKACESESTDTNTBHQ 240  
 |||||||  
 Db 181 ELGSDSSEDTYKATYCGVDOELQITPOGTRDEISLSAKKACESESTDTNTBHQ 240

QY 241 PSNNDLNTTEKRAAERHPEKYOGSSVNLHVEPCGTNTHASSLOHENSLLTKDRNVE 300  
 |||||||  
 Db 241 PSNNDLNTTEKRAAERHPEKYOGSSVNLHVEPCGTNTHASSLOHENSLLTKDRNVE 300

QY 301 KAEFCNKSKOPCLANSQNNRNAGSKETCNDRTSTTEKKVULNDPLCERKENKOKLPC 360  
 |||||||  
 Db 301 KAEFCNKSKOPCLANSQNNRNAGSKETCNDRTSTTEKKVULNDPLCERKENKOKLPC 360

QY 361 SENPRDTEVPWITLNNSTLOKYNEWFSRSDLELSDSDSHDESBNKAVADVLTLEVD 420  
 |||||||  
 Db 361 SENPRDTEVPWITLNNSTLOKYNEWFSRSDLELSDSDSHDESBNKAVADVLTLEVD 420

QY 421 EYSGSSEKIDLLASDPHALICKSERVHASKVESNEIDKJFGKTYRRKASLPLNLSHTEN 480  
 |||||||  
 Db 421 EYSGSSEKIDLLASDPHALICKSERVHASKVESNEIDKJFGKTYRRKASLPLNLSHTEN 480

QY 481 LIIGAFVTEPQIIOGRPLTNLKKRRRPTSLGHEHEDPIKADLAVOKTPEMINGNOTE 540  
 |||||||  
 Db 481 LIIGAFVTEPQIIOGRPLTNLKKRRRPTSLGHEHEDPIKADLAVOKTPEMINGNOTE 540

QY 541 ONGOVNNTNSGHEKKTGDSIONENKNPNPESLEKESAFTRKAPLSSISSNMELENI 600  
 |||||||  
 Db 541 ONGOVNNTNSGHEKKTGDSIONENKNPNPESLEKESAFTRKAPLSSISSNMELENI 600

QY 601 HNSKAPKKRRLKRSSTRHIALLELVSRNLSPNCTELOIDSCSSSEETIKKKRYNOMPV 660  
 |||||||  
 Db 601 HNSKAPKKRRLKRSSTRHIALLELVSRNLSPNCTELOIDSCSSSEETIKKKRYNOMPV 660

QY 661 RHSRNLQMLEGKEPRTGAKKSNKPNEQTSKRHSDDTPEELKLTINAPSPFKCSWTSLKE 720  
 |||||||  
 Db 661 RHSRNLQMLEGKEPRTGAKKSNKPNEQTSKRHSDDTPEELKLTINAPSPFKCSWTSLKE 720

QY 721 FVNPLPREEKEKETVYKVNNAEDPDMLSEERYLOTERSVSSSISILVPGTDVGTQ 780  
 |||||||  
 Db 721 FVNPLPREEKEKETVYKVNNAEDPDMLSEERYLOTERSVSSSISILVPGTDVGTQ 780

QY 781 ESISLLEVSTLGAKTEBENKCVSOCAAFENRKGILHGCXKONRNDTGFYRPLGHEVNH 840  
 |||||||  
 Db 781 ESISLLEVSTLGAKTEBENKCVSOCAAFENRKGILHGCXKONRNDTGFYRPLGHEVNH 840

QY 841 RETSIEMEESLDAQYLONTFKVSKROSFALEFNSPNAEECATFSAHSSLSLKQSPKVT 900  
 |||||||  
 Db 841 RETSIEMEESLDAQYLONTFKVSKROSFALEFNSPNAEECATFSAHSSLSLKQSPKVT 900

QY 901 FECECKEENOKNENIKIPVQVNTAGFPVYGOKDKPVDNAKSTIGSGRPLCSSSOPFG 960  
 |||||||  
 Db 901 FECECKEENOKNENIKIPVQVNTAGFPVYGOKDKPVDNAKSTIGSGRPLCSSSOPFG 960

QY 961 NETGLITPNKGLLONPYRIPPLPPIKSFVTKCKKNLLEBNFEHSMSPEREMGNITP 1020



QY 121 EVSIISMGYRNRARKLLOSEPENPSLOETSLVSQLSNLTGVTRTLTKORIQOKTSVYI 180  
 DB 121 EVSIISMGYRNRARKLLOSEPENPSLOETSLVSQLSNLTGVTRTLTKORIQOKTSVYI 180  
 QY 181 ELGSDSSEDTYKATYCVSGDELLQITPOGTIDEISLDSAKKACEFSEDTVNTENHQ 240  
 DB 181 ELGSDSSEDTYKATYCVSGDELLQITPOGTIDEISLDSAKKACEFSEDTVNTENHQ 240  
 QY 241 PSNNLNTTEKRAAEHPEKYGSSVSNLHVEPCGNTNASSLOHENSLLTLKDMNVE 300  
 DB 241 PSNNLNTTEKRAAEHPEKYGSSVSNLHVEPCGNTNASSLOHENSLLTLKDMNVE 300  
 QY 301 KAEFCNKSQOPGLARSQHNRMAGSKETCNDRTPTSEKTKVDLNADELCEKKNKKOLPC 360  
 DB 301 KAEFCNKSQOPGLARSQHNRMAGSKETCNDRTPTSEKTKVDLNADELCEKKNKKOLPC 360  
 QY 361 SENPDDTEDVPWITLNNSTOKVNEWFSRSDDELGSDSDHGESESNAAVADVLVNEVD 420  
 DB 361 SENPDDTEDVPWITLNNSTOKVNEWFSRSDDELGSDSDHGESESNAAVADVLVNEVD 420  
 QY 421 EYSGSSEKIDLASDPEHALICKSERVHKSVESENTEDIKFGTKYRKASLPULSHVTEN 480  
 DB 421 EYSGSSEKIDLASDPEHALICKSERVHKSVESENTEDIKFGTKYRKASLPULSHVTEN 480  
 QY 481 LIIGAFTVEPOJIOERPLTNKLRKRRTSGLHPEDFIKKADLAVOKTEPMINOGTNOTE 540  
 DB 481 LIIGAFTVEPOJIOERPLTNKLRKRRTSGLHPEDFIKKADLAVOKTEPMINOGTNOTE 540  
 QY 541 ONGQVWNTNSGHEKTKDSTIONENPNPISLEKESAFKTKAEPISSISMELELNI 600  
 DB 541 ONGQVWNTNSGHEKTKDSTIONENPNPISLEKESAFKTKAEPISSISMELELNI 600  
 QY 601 HNSKAPKRNLRKRSSTRIHLELVSNLSPNCTEJOIDSCSSSEETKRRKYNOMPY 660  
 DB 601 HNSKAPKRNLRKRSSTRIHLELVSNLSPNCTEJOIDSCSSSEETKRRKYNOMPY 660  
 QY 661 RHRNRLQIMEGKEPATGAKSKNPNEQTSKRHSDSTFPPELKLTNAGSFTKCSNTSELKE 720  
 DB 661 RHRNRLQIMEGKEPATGAKSKNPNEQTSKRHSDSTFPPELKLTNAGSFTKCSNTSELKE 720  
 QY 721 FVNPSLPREKEKEKLETVKVSNNAPEDPKDMLSGEVLQTERSVESSISLVPGTGYGRO 780  
 DB 721 FVNPSLPREKEKEKLETVKVSNNAPEDPKDMLSGEVLQTERSVESSISLVPGTGYGRO 780  
 QY 781 ESISLLEVSTLGAKTEPNKCVSQAFAENPKGLIHGCSKDNNDTEGFKYPLGHEVNS 840  
 DB 781 ESISLLEVSTLGAKTEPNKCVSQAFAENPKGLIHGCSKDNNDTEGFKYPLGHEVNS 840  
 QY 841 RETSIEMESELDAQYLONTFVYSKROSFALFSNPGNAEECATFSAHSGSLKKQSPKVT 900  
 DB 841 RETSIEMESELDAQYLONTFVYSKROSFALFSNPGNAEECATFSAHSGSLKKQSPKVT 900  
 QY 901 FECBCKEENOGKNEGNIKPVOTVNTAGPVYQKXKPYDNMKCSIKGSSRCLSSQFRG 960  
 DB 901 FECBCKEENOGKNEGNIKPVOTVNTAGPVYQKXKPYDNMKCSIKGSSRCLSSQFRG 960  
 QY 961 NETGLITPNKHGLONPYRIPLPIKSFVKTKCKKNLLEENFEHSMSPEREMENENP 1020  
 DB 961 NETGLITPNKHGLONPYRIPLPIKSFVKTKCKKNLLEENFEHSMSPEREMENENP 1020  
 QY 1021 STVSTISRNNTRENVFKEAASSNINEVGSSTNEVGSSINEIGSSDENIOAELGRNRPKL 1080  
 DB 1021 STVSTISRNNTRENVFKEAASSNINEVGSSTNEVGSSINEIGSSDENIOAELGRNRPKL 1080  
 QY 1081 NAMURLGLVLOPEVYQSLPGSNCKHPETIKOEYEEVYQVNNDESYLLSDMLEQMGSS 1140  
 DB 1081 NAMURLGLVLOPEVYQSLPGSNCKHPETIKOEYEEVYQVNNDESYLLSDMLEQMGSS 1140  
 QY 1141 HASQVCSETPDDLDDGEIKEDTSFAENDIKESAVFSKSVOKGELSRSPFTHTHLLAO 1200  
 DB 1141 HASQVCSETPDDLDDGEIKEDTSFAENDIKESAVFSKSVOKGELSRSPFTHTHLLAO 1200

QY 1201 GYRGAKKLESSEENENLSEDEELPCFOHLLFGKVNNIPOSOTRHSVTATECLSKNTEENL 1260  
 DB 1201 GYRGAKKLESSEENENLSEDEELPCFOHLLFGKVNNIPOSOTRHSVTATECLSKNTEENL 1260  
 QY 1261 LSLKNSLNDCSNOVLLAKASQEHHLSEETKGCASLFSQCSSELEDTANTNTODPFLIGS 1320  
 DB 1261 LSLKNSLNDCSNOVLLAKASQEHHLSEETKGCASLFSQCSSELEDTANTNTODPFLIGS 1320  
 QY 1321 SKOMRHQSSQSGSLDKELVSDDEERGTLKENNOEBSQMSNLGEAASGCESTVSSE 1380  
 DB 1321 SKOMRHQSSQSGSLDKELVSDDEERGTLKENNOEBSQMSNLGEAASGCESTVSSE 1380  
 QY 1381 DCSGSSQSDILITLTOORDTMQHNLIKLOEMAELAVLEQHGOSQPSNSYPSIISSSALE 1440  
 DB 1381 DCSGSSQSDILITLTOORDTMQHNLIKLOEMAELAVLEQHGOSQPSNSYPSIISSSALE 1440  
 QY 1441 DLRNPEOSTSEKAVLTSOKSSEYPISONPEGLSADKFEVSADSSSTKKKEGVERSSPSK 1500  
 DB 1441 DLRNPEOSTSEKAVLTSOKSSEYPISONPEGLSADKFEVSADSSSTKKKEGVERSSPSK 1500  
 QY 1501 CFSLDORWYMHSCSGSLQMRNYPQOEBELIKVYDVEEQULESGPHDLTETSYLUPRODLEG 1560  
 DB 1501 CFSLDORWYMHSCSGSLQMRNYPQOEBELIKVYDVEEQULESGPHDLTETSYLUPRODLEG 1560  
 QY 1561 TPYLESGISLFSDDPESDPESEDRAPEASRVNIPSTSAKLVQPKVAESAQSPAATHT 1620  
 DB 1561 TPYLESGISLFSDDPESDPESEDRAPEASRVNIPSTSAKLVQPKVAESAQSPAATHT 1620  
 QY 1621 DTAGYNAMEESVREPELTASTERYNKRMSNVVSGGLIPEEMLYKRAKHHTLNL 1680  
 DB 1621 DTAGYNAMEESVREPELTASTERYNKRMSNVVSGGLIPEEMLYKRAKHHTLNL 1680  
 QY 1681 TEETTHVAKTDAEFECERTLKYFLGIAGKVVVSYFWVTOSIKRKMILNHDPEVRGDV 1740  
 DB 1681 TEETTHVAKTDAEFECERTLKYFLGIAGKVVVSYFWVTOSIKRKMILNHDPEVRGDV 1740  
 QY 1741 VNGRNHQGFKRARESDRKIFRGLIICCGPFTNPTQOLEMWOLCGASVVKELSSFTL 1800  
 DB 1741 VNGRNHQGFKRARESDRKIFRGLIICCGPFTNPTQOLEMWOLCGASVVKELSSFTL 1800  
 QY 1801 GTGVHPITVVOPDAMTEONGFHAIGOMCEAPVYPRFWLDSVALYQCOELDTYVLIPOIPH 1860  
 DB 1801 GTGVHPITVVOPDAMTEONGFHAIGOMCEAPVYPRFWLDSVALYQCOELDTYVLIPOIPH 1860  
 QY 1861 SHY 1863  
 DB 1861 SHY 1863  
 RESULT 15  
 AAR81532  
 ID AAR81532 standard; Protein; 1863 AA.  
 XX  
 AC AAR81532;  
 DT 02-OCT-1996 (first entry)  
 XX  
 DE BRCA1 mutant from sample set MSK7542.  
 XX  
 KW Cancer therapy; breast and ovarian cancer predisposing gene; immunogen;  
 antibody production; germline alteration; probe; lesion neoplasia; human;  
 gene therapy; protein replacement therapy; protein mimetic; BRCA1.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 1852 /note= "T1852S"  
 FT  
 XX  
 PN MO9605306-A2.  
 XX  
 PD 22-FEB-1996.  
 XX





```

OY 1441 DLRNPEOSTSEKAVLTSOKSSEYPISONPEGLSADKFEVSADOSTSKNKEPGYERSSPSK 1500
    |||||
Db 1441 DLRNPEOSTSEKAVLTSOKSSEYPISONPEGLSADKFEVSADOSTSKNKEPGYERSSPSK 1500
OY 1501 CPSLDDRWYMHSCSGSLQNRNYPISOELIKVVDVEEQOLEESGPHDLTETSYPRODLEG 1560
    |||||
Db 1501 CPSLDDRWYMHSCSGSLQNRNYPISOELIKVVDVEEQOLEESGPHDLTETSYPRODLEG 1560
OY 1561 TPYLESGLTFSDDPESDPSSEDRAPESARVGNIPSSTSALKVPOLKVAESAOSPAAHHT 1620
    |||||
Db 1561 TPYLESGLTFSDDPESDPSSEDRAPESARVGNIPSSTSALKVPOLKVAESAOSPAAHHT 1620
OY 1621 DTAGYNAMEESYSREKPELTASTERYNKRMSWVSGLTPEEFMLVYKFAKHHITLTNLI 1680
    |||||
Db 1621 DTAGYNAMEESYSREKPELTASTERYNKRMSWVSGLTPEEFMLVYKFAKHHITLTNLI 1680
OY 1681 TEETHHVVMKTDAEFVCERTLKFFLGIAAGKVVSYFWVTOSIKERKMLNEHDFEVRGDV 1740
    |||||
Db 1681 TEETHHVVMKTDAEFVCERTLKFFLGIAAGKVVSYFWVTOSIKERKMLNEHDFEVRGDV 1740
OY 1741 VNGRHHOGPKRARESODRKIFRGLETCCYGPFTNMPTDLEMMVOLCGASVVKELSSFTL 1800
    |||||
Db 1741 VNGRHHOGPKRARESODRKIFRGLETCCYGPFTNMPTDLEMMVOLCGASVVKELSSFTL 1800
OY 1801 GTGVHPVVQPDAMTEDNGFHAIGOMCEAPVVTREMLDSVALYOCOEIDTYLIPQIPH 1860
    |||||
Db 1801 GTGVHPVVQPDAMTEDNGFHAIGOMCEAPVVTREMLDSVALYOCOEIDTYLIPQIPH 1860
OY 1861 SHY 1863
    |||
Db 1861 SHY 1863

```

Search completed: June 27, 2003, 10:22:32  
 Job time : 100 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: June 27, 2003, 10:20:56 ; Search time 45 Seconds  
(without alignments)  
4539.379 Million cell updates/sec

Title: US-09-734-672-4  
9649  
Perfect score: 1 MDLSALRVEEYVNIAMOK.....LYCCQELDTYLIPQIPSHY 1863  
Sequence:

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 424699 seqs, 109646833 residues

Total number of hits satisfying chosen parameters: 424699

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

Published\_Applications\_AA:\*  
1: /cgn2\_6/ptodata/1/pubppaa/US08\_NEW\_PUB.pep:\*  
2: /cgn2\_6/ptodata/1/pubppaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubppaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubppaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubppaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB.pep:\*  
8: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubppaa/US09\_NEW\_PUB.pep:\*  
10: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubppaa/US10\_NEW\_PUB.pep:\*  
12: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/1/pubppaa/US60\_NEW\_PUB.pep:\*  
14: /cgn2\_6/ptodata/1/pubppaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9649	100.0	1863	US-09-734-672-4	Sequence 4, Appli
2	9649	100.0	1863	US-09-982-828-6	Sequence 6, Appli
3	9635	99.9	1863	US-09-734-672-2	Sequence 2, Appli
4	9635	99.9	1863	US-09-734-672-6	Sequence 6, Appli
5	9635	99.9	1863	US-09-982-828-2	Sequence 2, Appli
6	9635	99.9	1863	US-09-982-828-4	Sequence 4, Appli
7	9627	99.8	1863	US-10-022-819-2	Sequence 2, Appli
8	357	3.7	2344	US-09-815-242-12713	Sequence 12713, A
9	341	3.5	6281	US-09-815-242-12996	Sequence 12996, A
10	333	3.5	3899	US-10-171-311-4	Sequence 4, Appli
11	333	3.5	3917	US-10-171-311-8	Sequence 8, Appli
12	329	3.4	3907	US-10-171-311-2	Sequence 2, Appli
13	329	3.4	3925	US-10-171-311-6	Sequence 6, Appli
14	328.5	3.4	2368	US-09-815-242-5635	Sequence 5635, Ap
15	328.5	3.4	2368	US-09-815-242-12389	Sequence 12389, A
16	314	3.3	1596	US-09-902-432-4	Sequence 4, Appli
17	312	3.2	1400	US-09-764-176-7	Sequence 7, Appli
18	312	3.2	2478	US-09-815-242-5816	Sequence 5816, Ap
19	312	3.2	2478	US-09-815-242-12967	Sequence 12967, A

20	308	3.2	2665	US-09-864-761-34248	Sequence 34248, A
21	306	3.2	3158	US-09-815-242-12611	Sequence 12611, A
22	299.5	3.1	2843	US-09-987-482-1	Sequence 1, Appli
23	298.5	3.1	2843	US-08-681-219-32	Sequence 32, Appli
24	289	3.0	1979	US-10-205-823-419	Sequence 419, App
25	288	3.0	2476	US-09-824-574-7	Sequence 7, Appli
26	286.5	3.0	1781	US-09-961-403-13	Sequence 13, Appli
27	286.5	3.0	1781	US-09-738-877-3	Sequence 3, Appli
28	285.5	3.0	2025	US-09-815-242-5703	Sequence 5703, Ap
29	282	2.9	50	US-09-998-667-15	Sequence 15, Appli
30	278	2.9	3256	US-09-919-039-21	Sequence 21, Appli
31	278	2.9	3256	US-09-919-172-98	Sequence 98, Appli
32	272	2.8	2437	US-09-815-242-5834	Sequence 5834, Ap
33	268	2.8	5795	US-09-815-242-12610	Sequence 12610, A
34	267.5	2.8	1664	US-09-978-343-2	Sequence 2, Appli
35	267.5	2.8	1664	US-09-964-858-1	Sequence 1, Appli
36	264	2.7	1639	US-10-087-464-10	Sequence 10, Appli
37	262.5	2.7	1312	US-10-071-179-29	Sequence 29, Appli
38	261	2.7	1285	US-09-982-091A-2	Sequence 2, Appli
39	261	2.7	1938	US-10-171-311-164	Sequence 164, App
40	261	2.7	1945	US-09-927-597-2	Sequence 2, Appli
41	259	2.7	2383	US-10-082-830-260	Sequence 260, App
42	258.5	2.7	1972	US-10-171-311-162	Sequence 162, App
43	258.5	2.7	1979	US-09-927-597-4	Sequence 4, Appli
44	258	2.7	1786	US-09-742-096-3	Sequence 3, Appli
45	256	2.7	1616	US-09-820-843A-16	Sequence 16, Appli

## ALIGNMENTS

RESULT 1  
US-09-734-672-4  
Sequence 4, Application US/09734672  
Publication No. US20020183268A1

## GENERAL INFORMATION:

APPLICANT: Murphy, Patricia D.

Allen, Antonette C.

Alvarez, Christopher P.

Critz, Brenda S.

Olson, Sheri J.

Schelter, Denise B.

Zeng, Bin

TITLE OF INVENTION: Coding Sequences of the Human

NUMBER OF SEQUENCES: 72

CORRESPONDENCE ADDRESSES:

ADDRESS: Morgan Lewis & Bockius LLP

STREET: 1111 Pennsylvania Ave., N.W.

CITY: Washington

STATE: District of Columbia

COUNTRY: USA

ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/734.672

FILING DATE: 03-Dec-2000

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/966,436

FILING DATE: 07-No. US20020183268A1-97

APPLICATION NUMBER: US 08/598,591

FILING DATE: 12-Feb-96

ATTORNEY/AGENT INFORMATION:

NAME: Michael S. Tuscan

REGISTRATION NUMBER: 43,210

REFERENCE/DOCKET NUMBER: 44921-5055-02-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-739-3000

TELEFAX: 202-739-3001

```

: INFORMATION FOR SEQ ID NO: 4:
:   SEQUENCE CHARACTERISTICS:
:     LENGTH: 1863 amino acids
:     TYPE: amino acid
:     STRANDEDNESS: No. US20020183268A1 Relevant
:     TOPOLOGY: No. US20020183268A1 Relevant
:     MOLECULE TYPE: protein
:     ORIGINAL SOURCE:
:       ORGANISM: Homo sapiens
:       STRAIN: BRCAl
:     POSITION IN GENOME:
:       CHROMOSOME/SEGMENT: 17
:       MAP POSITION: 17q21
:   SEQUENCE DESCRIPTION: SEQ ID NO: 4:
: US-09-734-672-4

Query Match      100.0%: Score 9649; DB 9; Length 1863;
Best Local Similarity 100.0%: Pred. No. 0;
Matches 1863; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDLSALRVEQVQVIVINAMOKIIECPICLDELKEPVTCKDHIFCKFCMKLKNOKKGPQ 60
DB 1 MDLSALRVEQVQVIVINAMOKIIECPICLDELKEPVTCKDHIFCKFCMKLKNOKKGPQ 60
QY 61 CPLCKNDITKRSLOESTRESQVLEELIKITICAFQDGTGLEVANSYNFAKKENNSPEHLK 120
DB 61 CPLCKNDITKRSLOESTRESQVLEELIKITICAFQDGTGLEVANSYNFAKKENNSPEHLK 120
QY 121 EVSIIOSMGYRRNARAKLLQSEPNPISLOETSVOISNIGTVTLTKKQIQOKTSVYI 180
DB 121 EVSIIOSMGYRRNARAKLLQSEPNPISLOETSVOISNIGTVTLTKKQIQOKTSVYI 180
QY 181 ELGSDSEDTYVKATCYSGVDOELLQITQGTREDEISLDSAKKACEFSETDVTNTEHHQ 240
DB 181 ELGSDSEDTYVKATCYSGVDOELLQITQGTREDEISLDSAKKACEFSETDVTNTEHHQ 240
QY 241 PENNDLNTTEKRAAEHPEKYQSSVSNLHVEPCGNTTAASSLQIHENSSLLTKDRMNV 300
DB 241 PENNDLNTTEKRAAEHPEKYQSSVSNLHVEPCGNTTAASSLQIHENSSLLTKDRMNV 300
QY 301 KAEFCCKSKOPGLARSOHNRWAGSKETCDNDRTPSTIEKRYVDLADDLCEKKNKOLPC 360
DB 301 KAEFCCKSKOPGLARSOHNRWAGSKETCDNDRTPSTIEKRYVDLADDLCEKKNKOLPC 360
QY 361 SENPRDEDPVITLSSIQKVNEFMSRDELLGSDSHDGESESNAKVADVDLVNEVD 420
DB 361 SENPRDEDPVITLSSIQKVNEFMSRDELLGSDSHDGESESNAKVADVDLVNEVD 420
QY 421 EYSGSSEKIDLLASDPHEALICKSEYVHSHSVESNIEDKIIFGKTYRKKASLPMLSHVTEN 480
DB 421 EYSGSSEKIDLLASDPHEALICKSEYVHSHSVESNIEDKIIFGKTYRKKASLPMLSHVTEN 480
QY 481 LITGAFTVEPQIIQERPLTNKLRKRKPTSGLHPEDFIKKADLAIVKTPMIMQGNQHE 540
DB 481 LITGAFTVEPQIIQERPLTNKLRKRKPTSGLHPEDFIKKADLAIVKTPMIMQGNQHE 540
QY 541 QNGQVWNITNSGHENKTGDSIQNEKNPNPISLESLESAFKTAEIISISSISIMLELNI 600
DB 541 QNGQVWNITNSGHENKTGDSIQNEKNPNPISLESLESAFKTAEIISISSISIMLELNI 600
QY 601 HNSKAPKRNRLRRKSSSTRIHALELVYRNLSPPNCTELOIDSCSSSEELIKKKKYQMRY 660
DB 601 HNSKAPKRNRLRRKSSSTRIHALELVYRNLSPPNCTELOIDSCSSSEELIKKKKYQMRY 660
QY 661 RHSRLQIMEKREPATGAKKSNKPNOTSKRHDSDPEFPELKLITNAGSTKCSNTSELKE 720
DB 661 RHSRLQIMEKREPATGAKKSNKPNOTSKRHDSDPEFPELKLITNAGSTKCSNTSELKE 720
QY 721 FVNPSLPREKEDEKLETYVSNNAEDPKDMLSGEFTVLOTERSVESSISLVPGTGYGTQ 780
DB 721 FVNPSLPREKEDEKLETYVSNNAEDPKDMLSGEFTVLOTERSVESSISLVPGTGYGTQ 780
QY 781 ESIISLEVTGAKTEPNKCVSQAFAENPKGLIHGCSKDNKRNDETEGFKYPLGHEVNS 840
```

```

DB 781 ESIISLEVTGAKTEPNKCVSQAFAENPKGLIHGCSKDNKRNDETEGFKYPLGHEVNS 840
QY 841 RETSIMESEELDAOVLQNTFVKRSQSFALSNNGMAEEECATTSANSGSLKQSPRYT 900
DB 841 RETSIMESEELDAOVLQNTFVKRSQSFALSNNGMAEEECATTSANSGSLKQSPRYT 900
QY 901 FECEOKEENQKNESNIRKPVQVNTIAGFPVVGOKDPVDNAKCSIKGSRFCLSSQPRG 960
DB 901 FECEOKEENQKNESNIRKPVQVNTIAGFPVVGOKDPVDNAKCSIKGSRFCLSSQPRG 960
QY 961 NETGLITPNKHGLQNPYRIPLPLPIKSFVTKCKKNLLEENFEESHMSPEREMGNENIP 1020
DB 961 NETGLITPNKHGLQNPYRIPLPLPIKSFVTKCKKNLLEENFEESHMSPEREMGNENIP 1020
QY 1021 STVSTISRNIRPNVKEASSNINNGVSTNEDVSSINELGSSDENITQAEIGRRRGPKL 1080
DB 1021 STVSTISRNIRPNVKEASSNINNGVSTNEDVSSINELGSSDENITQAEIGRRRGPKL 1080
QY 1081 NAMLRLGVLOPEVYKQSLGPNCKHPEIKKQYEVEVQTVNTDFSPYLLSDNLEQPMGSS 1140
DB 1081 NAMLRLGVLOPEVYKQSLGPNCKHPEIKKQYEVEVQTVNTDFSPYLLSDNLEQPMGSS 1140
QY 1141 HASQVCEPDPDLDLDGEIKEDTSPFAENDIKESSAVFSKSYQKGLSRSPPFTHTHLAQ 1200
DB 1141 HASQVCEPDPDLDLDGEIKEDTSPFAENDIKESSAVFSKSYQKGLSRSPPFTHTHLAQ 1200
QY 1201 GYRGAKKLESSEFENLSEDEELPCQHLLPCKVNNINISQSRHSTVATECLSKNTEENL 1260
DB 1201 GYRGAKKLESSEFENLSEDEELPCQHLLPCKVNNINISQSRHSTVATECLSKNTEENL 1260
QY 1261 LSLKNSLNDCSNOVILAKASQEHLSSEETKCSASFSSQCSLELDTANTNTQDPFLGS 1320
DB 1261 LSLKNSLNDCSNOVILAKASQEHLSSEETKCSASFSSQCSLELDTANTNTQDPFLGS 1320
QY 1321 SKOMRHOSQGVGLSKDELVSDDERGTGLEENNOBEOBMSNUGERASGCESETSYSE 1380
DB 1321 SKOMRHOSQGVGLSKDELVSDDERGTGLEENNOBEOBMSNUGERASGCESETSYSE 1380
QY 1381 DCSGLSSQSDILITTOORDTMQNLKILQOEMALEVLEOHGSOPTSNSTPSTISSSAL 1440
DB 1381 DCSGLSSQSDILITTOORDTMQNLKILQOEMALEVLEOHGSOPTSNSTPSTISSSAL 1440
QY 1441 DLNRPQSTSEKAVILTSQKSEYPISQNPEGLSADKFEVSADSTSKNKEPVERSSPSK 1500
DB 1441 DLNRPQSTSEKAVILTSQKSEYPISQNPEGLSADKFEVSADSTSKNKEPVERSSPSK 1500
QY 1501 CFSLDRLWYHSCSGSLQNRNTPSOBELIKVYDVEEQULEBSGPHDLITETSYLPRODLEG 1560
DB 1501 CFSLDRLWYHSCSGSLQNRNTPSOBELIKVYDVEEQULEBSGPHDLITETSYLPRODLEG 1560
QY 1561 TPYLESGISLFDSDPSPSEDRAPESARVGIIPSTSLAKVPOLKVESASQSPAATHT 1620
DB 1561 TPYLESGISLFDSDPSPSEDRAPESARVGIIPSTSLAKVPOLKVESASQSPAATHT 1620
QY 1621 DTAGYNAMEESVSRKEPDLTASTERYNKRMSWVSGLTPPEERMLYKFAKKHHTLTNLI 1680
DB 1621 DTAGYNAMEESVSRKEPDLTASTERYNKRMSWVSGLTPPEERMLYKFAKKHHTLTNLI 1680
QY 1681 TEETHYVAKTDAEYECERTLKYFLIGIAGKVVSYFWVTOSIKERKMLNHDPEVNGDV 1740
DB 1681 TEETHYVAKTDAEYECERTLKYFLIGIAGKVVSYFWVTOSIKERKMLNHDPEVNGDV 1740
QY 1741 VGRNHOQFKRARESDOKIFRGLEICCYGPTNPTQOLBEMVQOLGASVYKELSFLL 1800
DB 1741 VGRNHOQFKRARESDOKIFRGLEICCYGPTNPTQOLBEMVQOLGASVYKELSFLL 1800
QY 1801 GTGVHPPIVVQDPANTEDNGFHAIGOMCEAPVTRFWVLDSVALYOCQELDTYLIPOIPH 1860
DB 1801 GTGVHPPIVVQDPANTEDNGFHAIGOMCEAPVTRFWVLDSVALYOCQELDTYLIPOIPH 1860
QY 1861 SHY 1863
DB 1861 SHY 1863
```

Db 1861 SHY 1863

RESULT 2  
US-09-982-828-6  
Sequence 6, Application US/09982828  
Publication No. US20030022184A1  
GENERAL INFORMATION:  
APPLICANT: Murphy, Patricia D.  
Allen, Antonette C.  
Alvares, Christopher P.  
Critz, Brenda S.  
Olson, Sheri J.  
Thurber, Denise  
Zeng, Bin

TITLE OF INVENTION: Coding Sequences of the Human  
BRCA1 Gene

NUMBER OF SEQUENCES: 72  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morgan Lewis & Bockius LLP  
STREET: 1111 Pennsylvania Avenue N. W.  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20004

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/982,828  
FILING DATE: 22-Oct-2001  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/074,453  
FILING DATE: 1998-05-06  
APPLICATION NUMBER: US 08/798,691  
FILING DATE: 1997-02-12  
APPLICATION NUMBER: US 08/598,591  
FILING DATE: 1996-02-12

ATTORNEY/AGENT INFORMATION:  
NAME: Michael S. Tuscan  
REGISTRATION NUMBER: 43,210  
REFERENCE/DOCKET NUMBER: 44921-5053-01-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-739-3000  
TELEFAX: 202-739-3001

INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1863 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
STRAIN: BRCA1 (cm13)  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: 17  
MAP POSITION: 17q21  
SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
US-09-982-828-6

Query Match 100.0%; Score 9649; DB 9; Length 1863;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1863; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDLSALRVEEVQVNIAMOKILCPICLETIKPEVSTKCDHIFCKFCMLKILNQKGPQ 60  
Db 1 MDLSALRVEEVQVNIAMOKILCPICLETIKPEVSTKCDHIFCKFCMLKILNQKGPQ 60

QY 61 CPLCKNDITKRSLOESTRFSQVLEELIKICAFOLDTGLEAVANSYNFAKKENNSPEHLKD 120

Db 61 CPLCKNDITKRSLOESTRFSQVLEELIKICAFOLDTGLEAVANSYNFAKKENNSPEHLKD 120

QY 121 EVSTIOSMGYRNRAKRLJOSEPEPNSLOETSLVSQVSLNCTVRLTRKRIQOKTSVYI 180  
Db 121 EVSTIOSMGYRNRAKRLJOSEPEPNSLOETSLVSQVSLNCTVRLTRKRIQOKTSVYI 180

QY 181 ELGSDSSEEDVNNKATYCSVGDDELQITPGTRDEISLDSAKRAACEFSETDVTNTEHHQ 240  
Db 181 ELGSDSSEEDVNNKATYCSVGDDELQITPGTRDEISLDSAKRAACEFSETDVTNTEHHQ 240

QY 241 PSNNDLNTTEKRAAERHPEKYQSSVSNLHVEPCGTNTAASLOHENSLLLTDRMNV 300  
Db 241 PSNNDLNTTEKRAAERHPEKYQSSVSNLHVEPCGTNTAASLOHENSLLLTDRMNV 300

QY 301 KAEFCNKSQPGIARSGHNWAGSKFCNDRPTSTKRYDLVADPLCEKKNKOKLPC 360  
Db 301 KAEFCNKSQPGIARSGHNWAGSKFCNDRPTSTKRYDLVADPLCEKKNKOKLPC 360

QY 361 SENPROTEDVPMITLINSIQVNEWFGRSDELGSDSDHGESESNKAVADVLDVNEVD 420  
Db 361 SENPROTEDVPMITLINSIQVNEWFGRSDELGSDSDHGESESNKAVADVLDVNEVD 420

QY 421 EYSGSSEKIDLLASDPHEALICKSERVHKSVSBNIEDKIFGKTYRRKASLPNLSHVTEN 480  
Db 421 EYSGSSEKIDLLASDPHEALICKSERVHKSVSBNIEDKIFGKTYRRKASLPNLSHVTEN 480

QY 481 LIIGAFTEPQIIIOERLTNKLKRRRPTSGIHPEDEIKRADLAVOCTPPEMINGNOTE 540  
Db 481 LIIGAFTEPQIIIOERLTNKLKRRRPTSGIHPEDEIKRADLAVOCTPPEMINGNOTE 540

QY 541 QNCOVNMINTSGHENKTGDSIONEKNPNIESLEKESAKRTAEPISISSINMTELENI 600  
Db 541 QNCOVNMINTSGHENKTGDSIONEKNPNIESLEKESAKRTAEPISISSINMTELENI 600

QY 601 HNSKAPKKNRLRRKSSSTRHIALELVVSRNLSPNCTEIQIDSCSSSEETIKKKYQMVP 660  
Db 601 HNSKAPKKNRLRRKSSSTRHIALELVVSRNLSPNCTEIQIDSCSSSEETIKKKYQMVP 660

QY 661 RHSRNQLMGKCPATGAKKSNKPNEDTSKRHSDTFPELKLINAPSPFKGNTSELKE 720  
Db 661 RHSRNQLMGKCPATGAKKSNKPNEDTSKRHSDTFPELKLINAPSPFKGNTSELKE 720

QY 721 FVNPSPPREKEKLELVKVSNNAEDEPKDMLSGERYLOTREYESSISLVPGTDYGTQ 780  
Db 721 FVNPSPPREKEKLELVKVSNNAEDEPKDMLSGERYLOTREYESSISLVPGTDYGTQ 780

QY 781 ESISLLEVSTLGAKTEPNKCVSQAFAFENPKGLIHGCSKDNBNDETFGRYPLGHEVNH 840  
Db 781 ESISLLEVSTLGAKTEPNKCVSQAFAFENPKGLIHGCSKDNBNDETFGRYPLGHEVNH 840

QY 841 RETSIEMESELDAYLIONTFKVSKROSFALEFNPAGAEDECATFSAHSGSLKKOSKVT 900  
Db 841 RETSIEMESELDAYLIONTFKVSKROSFALEFNPAGAEDECATFSAHSGSLKKOSKVT 900

QY 901 FECEOKKEENGKNESNKPQVNTNITAGFPVVOCKDPVYNACSLIKGSRFCLSSOFRG 960  
Db 901 FECEOKKEENGKNESNKPQVNTNITAGFPVVOCKDPVYNACSLIKGSRFCLSSOFRG 960

QY 961 NETGLITPNKHGILLQNPYRIPLPEIKSFVTKCKKNLLEENFEHSMSPEREMGENIP 1020  
Db 961 NETGLITPNKHGILLQNPYRIPLPEIKSFVTKCKKNLLEENFEHSMSPEREMGENIP 1020

QY 1021 STVSTISRRNIRNVRKEASSNINEVGSSSTNEVGSSINTEGSDDEITQAFGRNGPKL 1080  
Db 1021 STVSTISRRNIRNVRKEASSNINEVGSSSTNEVGSSINTEGSDDEITQAFGRNGPKL 1080

QY 1081 NAMLRGLVLOPEYKQSLPSCNCKHPEIKKQEEVVOVNTPEFSYLIISDNLEQMGSS 1140  
Db 1081 NAMLRGLVLOPEYKQSLPSCNCKHPEIKKQEEVVOVNTPEFSYLIISDNLEQMGSS 1140

QY 1141 HASOVSETPDDLLDDEIKEDTSAENDIKESSAVFSKVOGELSRSPPTTHLQ 1200

Db 1141 HASQVCSETPDDLDDGEIKEDTSEFANDIKESSAAVFSKSVQKGLSRSPFTHTHLAQ 1200  
 QY 1201 GYRGAKKLESSEENLSEDEELPCQHLFLGKVNINIPSOSTRSHVATECLSKNTEENL 1260  
 Db 1201 GYRRAKKLESSEENLSEDEELPCQHLFLGKVNINIPSOSTRSHVATECLSKNTEENL 1260  
 QY 1261 LSKNSLNDSCNOVILAKASQEHHLSEETKCSASLFSQCSSELEDLTANTNTODPPLIGS 1320  
 Db 1261 LSKNSLNDSCNOVILAKASQEHHLSEETKCSASLFSQCSSELEDLTANTNTODPPLIGS 1320  
 QY 1321 SKOMHOSQGVGLSDKELVSDDEERGTLGLENNOEODSMNDLGEASGCESETSVSE 1380  
 Db 1321 SKOMHOSQGVGLSDKELVSDDEERGTLGLENNOEODSMNDLGEASGCESETSVSE 1380  
 QY 1381 DCSGLSSQSDILITTOORDTMHNLKLOQEMAELEAVLEBQHSQSPNSYPSIISDSALE 1440  
 Db 1381 DCSGLSSQSDILITTOORDTMHNLKLOQEMAELEAVLEBQHSQSPNSYPSIISDSALE 1440  
 QY 1441 DLRNPEQSTSEKAVLTSQKSEYPISQNPGLSADKFEVSADSTSKNKEPVERSSPSK 1500  
 Db 1441 DLRNPEQSTSEKAVLTSQKSEYPISQNPGLSADKFEVSADSTSKNKEPVERSSPSK 1500  
 QY 1501 CPSLDDRWYMHSCSGLONRNPPOBELIKVYDVEEQOLEESGPHDLTETSTYLPDRLDG 1560  
 Db 1501 CPSLDDRWYMHSCSGLONRNPPOBELIKVYDVEEQOLEESGPHDLTETSTYLPDRLDG 1560  
 QY 1561 TPYLESGISLFDDEPDESDERAPESARVGNIPSTSAKLVQKVAESAQSPAAAHHT 1620  
 Db 1561 TPYLESGISLFDDEPDESDERAPESARVGNIPSTSAKLVQKVAESAQSPAAAHHT 1620  
 QY 1621 DTAGYNAMEESVSRKPELTASTERYNKRMSNVVSGLTPEEFMLVYKFAKHHTITLNL 1680  
 Db 1621 DTAGYNAMEESVSRKPELTASTERYNKRMSNVVSGLTPEEFMLVYKFAKHHTITLNL 1680  
 QY 1681 TEETHHVYKTTAEFCERTILKYFLGIAGKWWVSFWYVQSTKEKKNLNEHFEVGRGV 1740  
 Db 1681 TEETHHVYKTTAEFCERTILKYFLGIAGKWWVSFWYVQSTKEKKNLNEHFEVGRGV 1740  
 QY 1741 VNGRNHOGPKRARESDRKIFRGLGICCGPFTNMPTDLEMMVQCGASVYKELSSFTL 1800  
 Db 1741 VNGRNHOGPKRARESDRKIFRGLGICCGPFTNMPTDLEMMVQCGASVYKELSSFTL 1800  
 QY 1801 GTGVPIYVOPDANTEGNGFHAIGOMCAPVYVTRWVLDVALVOCQELDTYLLIPQIPH 1860  
 Db 1801 GTGVPIYVOPDANTEGNGFHAIGOMCAPVYVTRWVLDVALVOCQELDTYLLIPQIPH 1860  
 QY 1861 SHY 1863  
 Db 1861 SHY 1863

RESULT 3  
 US-09-734-672-2  
 : Sequence 2, Application us/09734672  
 : Publication No. US20020183268A1  
 : GENERAL INFORMATION:

APPLICANT: Murphy, Patricia D.

Allen, Antoinette C.

Alvares, Christopher P.

Critz, Brenda S.

Olson, Sheri J.

Schelter, Denise B.

Zeng, Bin

TITLE OF INVENTION: Coding Sequences of the Human

BCR1 Gene

NUMBER OF SEQUENCES: 72

CORRESPONDENCE ADDRESS:

ADDRESSSEE: Morgan Lewis & Bockius LLP

STREET: 1111 Pennsylvania Ave., N.W.

CITY: Washington

STATE: District of Columbia

COUNTRY: USA

ZIP: 20004

? COMPUTER READABLE FORM:  
 ? MEDIUM TYPE: Floppy disk  
 ? COMPUTER: IBM PC compatible  
 ? OPERATING SYSTEM: PC-DOS/MS-DOS  
 ? SOFTWARE: Patentin Release #1.0, Version #1.30  
 ? CURRENT APPLICATION DATA:  
 ? APPLICATION NUMBER: US/09/734,672  
 ? FILING DATE: 03-Dec-2000  
 ? PRIOR APPLICATION DATA:  
 ? APPLICATION NUMBER: US 08/966,436  
 ? FILING DATE: 07-Nov-96  
 ? APPLICATION NUMBER: US 08/598,591  
 ? FILING DATE: 12-Feb-96  
 ? ATTORNEY/AGENT INFORMATION:  
 ? NAME: Michael S. Tuscan  
 ? REGISTRATION NUMBER: 43,210  
 ? REFERENCE/DOCKET NUMBER: 44921-5055-02-US  
 ? TELECOMMUNICATION INFORMATION:  
 ? TELEPHONE: 202-739-3001  
 ? TELEFAX: 202-739-3000  
 ? INFORMATION FOR SEQ ID NO: 2:  
 ? SEQUENCE CHARACTERISTICS:  
 ? LENGTH: 1863 amino acids  
 ? TYPE: amino acid  
 ? STRANDEDNESS: No. US20020183268A1 Relevant  
 ? TOPOLOGY: No. US20020183268A1 Relevant  
 ? MOLECULE TYPE: protein  
 ? ORIGINAL SOURCE:  
 ? ORGANISM: Homo sapiens  
 ? STRAIN: BCR1  
 ? POSITION IN GENOME:  
 ? CHROMOSOME/SEGMENT: 17  
 ? MAP POSITION: 17q21  
 ? SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
 ? US-09-734-672-2  
 ?  
 ? Query Match 99.9%; Score 9635; DB 9; Length 1863;  
 ? Best Local Similarity 99.8%; Pred. No. 0;  
 ? Matches 1860; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 ?  
 QY 1 MSLAARVEONVINAMOKILCEPICLELKEPSTYCDHIFCFCKLKLNNKKRGSQ 60  
 Db 1 MSLAARVEONVINAMOKILCEPICLELKEPSTYCDHIFCFCKLKLNNKKRGSQ 60  
 QY 61 CPLCKNDIKRSLQESTRFSQVLEELKICAFQDGTGLEYANSYNFAKKENNSPEHLKD 120  
 Db 61 CPLCKNDIKRSLQESTRFSQVLEELKICAFQDGTGLEYANSYNFAKKENNSPEHLKD 120  
 QY 121 EVSIIOGKGRNRAKRLQSEPNPSLQETSLSVOLNMGVYRILTKRIQIOPKTSYI 180  
 Db 121 EVSIIOGKGRNRAKRLQSEPNPSLQETSLSVOLNMGVYRILTKRIQIOPKTSYI 180  
 QY 181 ELGSDSEEDTVNKATYCSVGOELQITPQGRDEISLDSAKKAACESEEDVNTHEHQ 240  
 Db 181 ELGSDSEEDTVNKATYCSVGOELQITPQGRDEISLDSAKKAACESEEDVNTHEHQ 240  
 QY 241 PSNNDLNTEKRAAERHEKYOGSSVSNLHVEPCGNTNHAASLOHENSLLITDRNVE 300  
 Db 241 PSNNDLNTEKRAAERHEKYOGSSVSNLHVEPCGNTNHAASLOHENSLLITDRNVE 300  
 QY 301 KAEFCNKSQOPGLARSQHNRMAGSKETCNDRTSTTEKKVLDLNDPLCEKRNKKOKLPC 360  
 Db 301 KAEFCNKSQOPGLARSQHNRMAGSKETCNDRTSTTEKKVLDLNDPLCEKRNKKOKLPC 360  
 QY 361 SENPRDTEVPWITLNNSTIQKVENEFSSDELIGSDSDHDESSNMKVAADVLDVLEVD 420  
 Db 361 SENPRDTEVPWITLNNSTIQKVENEFSSDELIGSDSDHDESSNMKVAADVLDVLEVD 420  
 QY 421 EYSGSSEKIDLASPHALICKSERVASKSVESNIEBKIGKTYRKKASLPNLSHTEN 480  
 Db 421 EYSGSSEKIDLASPHALICKSERVASKSVESNIEBKIGKTYRKKASLPNLSHTEN 480  
 QY 481 LIIGAFVTEPOLIOERPLTNKLRKRRTSGLHPEDITKADLAVOKTPEMINOGTNOTE 540

Db 481 LIIAGVTEPOLIOERPLTKLKRKRRPTSGLHDEDFIKKADLAVOTKPTMINOQTOTE 540  
 QY 541 ONQOVNNTSGHENTKGDSTIONENKPNPIESLEKSAFKTAKEPSSISMETELNI 600  
 Db 541 ONQOVNNTSGHENTKGDSTIONENKPNPIESLEKSAFKTAKEPSSISMETELNI 600  
 QY 601 HNSAKRKNRLRRKSSSTRIHALELVYRNLSPPNCTELOIDSCSSSEETKRRKKNQMPV 660  
 Db 601 HNSAKRKNRLRRKSSSTRIHALELVYRNLSPPNCTELOIDSCSSSEETKRRKKNQMPV 660  
 QY 661 RHNRNQLMEGKPAAGAKSKNPNQTSKRHSDTFPELKLITNAPGSKNTSELKE 720  
 Db 661 RHNRNQLMEGKPAAGAKSKNPNQTSKRHSDTFPELKLITNAPGSKNTSELKE 720  
 QY 721 FVNPSPREKEKLEFVKNVSNNAEDPKDMLSGEVLQTERSVESSTISLVGTDTGTQ 780  
 Db 721 FVNPSPREKEKLEFVKNVSNNAEDPKDMLSGEVLQTERSVESSTISLVGTDTGTQ 780  
 QY 781 ESISLLEVSTLGRKAKTEPNKCVSQAFAFENPKGLHGCSKDNNDTEGEFYPGLGHEVNH 840  
 Db 781 ESISLLEVSTLGRKAKTEPNKCVSQAFAFENPKGLHGCSKDNNDTEGEFYPGLGHEVNH 840  
 QY 841 RETSIMESSELDQYLQNTFKVSKRQSFALSPNPGNAEECATFSAHSGSLKKQSPKVT 900  
 Db 841 RETSIMESSELDQYLQNTFKVSKRQSFALSPNPGNAEECATFSAHSGSLKKQSPKVT 900  
 QY 901 FECEOKEENQKKNESNIRKPVQVNTAGFPVVGOKKPVNACSTIKGSRFLSSQFRG 960  
 Db 901 FECEOKEENQKKNESNIRKPVQVNTAGFPVVGOKKPVNACSTIKGSRFLSSQFRG 960  
 QY 961 NETGLTPNKHGLQNPYRIPPLFPKISFVYTKCKKNLEENFEHSMSPREMGENIP 1020  
 Db 961 NETGLTPNKHGLQNPYRIPPLFPKISFVYTKCKKNLEENFEHSMSPREMGENIP 1020  
 QY 1021 STVSTISRNINRENFKEASSNINEVGSSTNEVSSINEIGSSDENIQALGRNGPKL 1080  
 Db 1021 STVSTISRNINRENFKEASSNINEVGSSTNEVSSINEIGSSDENIQALGRNGPKL 1080  
 QY 1081 NAMLRGLVLOPEYKQSLPGSNCKHPEIKKOEVEVVOYVNTDFSPLLSDNLEOPGSS 1140  
 Db 1081 NAMLRGLVLOPEYKQSLPGSNCKHPEIKKOEVEVVOYVNTDFSPLLSDNLEOPGSS 1140  
 QY 1141 HASQVSETPDDLDDGEIKEDTSFAENDIKESSAVFSKSVQKELSRSPPTHTHLAQ 1200  
 Db 1141 HASQVSETPDDLDDGEIKEDTSFAENDIKESSAVFSKSVQKELSRSPPTHTHLAQ 1200  
 QY 1201 GYRGRGAKKLESSEENLSSDEDELPCFQHLFGKVNIPQSQSTRHSTVATECLSKNTEENL 1260  
 Db 1201 GYRGRGAKKLESSEENLSSDEDELPCFQHLFGKVNIPQSQSTRHSTVATECLSKNTEENL 1260  
 QY 1261 LSLKNSLNDSCNOYLAKASQEHNLSEETKCSASLSSQCELEDTLANTNODPLIGS 1320  
 Db 1261 LSLKNSLNDSCNOYLAKASQEHNLSEETKCSASLSSQCELEDTLANTNODPLIGS 1320  
 QY 1321 SKQMRHQSQGVGLSDKELVSDDEERGTLGEENNOEQSMDSNLGEAASGCESETSVSE 1380  
 Db 1321 SKQMRHQSQGVGLSDKELVSDDEERGTLGEENNOEQSMDSNLGEAASGCESETSVSE 1380  
 QY 1381 DCSGLSSQSDILTTQORDTMQNLKLOQEMALEVLLEHGSQPSNSYPSITSDSALE 1440  
 Db 1381 DCSGLSSQSDILTTQORDTMQNLKLOQEMALEVLLEHGSQPSNSYPSITSDSALE 1440  
 QY 1441 DLNPNQOSTSEKAVLTSQKSESEYPISONPEGLSADFEVYASDSTSKNKEPGEVRSPPK 1500  
 Db 1441 DLNPNQOSTSEKAVLTSQKSESEYPISONPEGLSADFEVYASDSTSKNKEPGEVRSPPK 1500  
 QY 1501 CPSLDDRWYMHSCSGSLQNNRNPSEQLKVVYDVEEQLEESGPHULTSTYLPRLDLEG 1560  
 Db 1501 CPSLDDRWYMHSCSGSLQNNRNPSEQLKVVYDVEEQLEESGPHULTSTYLPRLDLEG 1560  
 QY 1561 TPYLEGISTLFSDDPSPSEDAAPESARVGNTPSSTSAKYNQLVYASQSPAAAHNTT 1620  
 Db 1561 TPYLEGISTLFSDDPSPSEDAAPESARVGNTPSSTSAKYNQLVYASQSPAAAHNTT 1620

Db 1561 TPYLEGISTLFSDDPSPSEDAAPESARVGNTPSSTSAKYNQLVYASQSPAAAHNTT 1620  
 QY 1621 DTAGYVNAEESVSRKREKELASTERVNRKMSMVVSGLTPEEFMLVYFARKHHTLTNLI 1680  
 Db 1621 DTAGYVNAEESVSRKREKELASTERVNRKMSMVVSGLTPEEFMLVYFARKHHTLTNLI 1680  
 QY 1681 TEETTHVAKTDAEFVCEKTLKTYFLGIAGKVVVSYFWVTSIKERKMLNHEFEVNGDV 1740  
 Db 1681 TEETTHVAKTDAEFVCEKTLKTYFLGIAGKVVVSYFWVTSIKERKMLNHEFEVNGDV 1740  
 QY 1741 VNGRNHOGPKRARESDQRKIFRGLIETCCYGFPTNMPDQLEMMVQLGASVVELSFTL 1800  
 Db 1741 VNGRNHOGPKRARESDQRKIFRGLIETCCYGFPTNMPDQLEMMVQLGASVVELSFTL 1800  
 QY 1801 GTGVHPVYVQPPAWTFEDNGFHAIGOMCEAPVYTRVWVLSVALYOCQELDTYLIPQIPH 1860  
 Db 1801 GTGVHPVYVQPPAWTFEDNGFHAIGOMCEAPVYTRVWVLSVALYOCQELDTYLIPQIPH 1860  
 QY 1861 SHY 1863  
 Db 1861 SHY 1863

RESULT 4  
 US-09-734-672-6  
 : Sequence 6, Application US/09734672  
 : Publication No. US20020183268A1  
 :  
 : GENERAL INFORMATION:  
 : APPLICANT: Murphy, Patricia D.  
 : Allen, Antonette C.  
 : Alvarez, Christopher P.  
 : Critz, Brenda S.  
 : Olson, Sheri J.  
 : Schelter, Denise B.  
 : Zeng, Bin  
 :  
 : TITLE OF INVENTION: Coding Sequences of the Human  
 : BRCA1 Gene  
 :  
 : NUMBER OF SEQUENCES: 72  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Morgan Lewis & Bockius LLP  
 : STREET: 1111 Pennsylvania Ave., N.W.  
 : CITY: Washington  
 : STATE: District of Columbia  
 : COUNTRY: USA  
 : ZIP: 20004  
 :  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: Patent In Release #1.0, Version #1.30  
 :  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/09/734,672  
 : FILING DATE: 03-Dec-2000  
 :  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: US 08/966,436  
 : FILING DATE: 07-No. US20020183268A1-97  
 : APPLICATION NUMBER: US 08/598,591  
 : FILING DATE: 12-Feb-96  
 :  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: Michael S. Tuscan  
 : REGISTRATION NUMBER: 43,210  
 : REFERENCE/DOCKET NUMBER: 44921-5055-02-US  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: 202-739-3000  
 : TELEFAX: 202-739-3001  
 :  
 : INFORMATION FOR SEQ ID NO: 6:  
 :  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 1863 amino acids  
 : TYPE: amino acid  
 : STRANDEDNESS: No. US20020183268A1 Relevant  
 : TOPOLOGY: No. US20020183268A1 Relevant  
 : MOLECULE TYPE: protein  
 :  
 : ORIGINAL SOURCE:



ORGANISM: Homo sapiens  
STRAIN: BRCA1  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: 17  
MAP POSITION: 17q21  
SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
US-09-734-672-6

Query Match 99.9% Score 9635; DB 9; Length 1863;  
Best Local Similarity 99.8% Pred. 0;  
Matches 1860; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDLSALRYEYVONVINAQKILECPICLDELKEPVSTKCDHIFCKFCMLKLLNOKKGPQ 60  
DB 1 MDLSALRYEYVONVINAQKILECPICLDELKEPVSTKCDHIFCKFCMLKLLNOKKGPQ 60  
QY 61 CPLCNDITKRSLOSTRSOLVLEELKIIICAFOLDTGLEYANSYNPAKKENSPHLD 120  
DB 61 CPLCNDITKRSLOSTRSOLVLEELKIIICAFOLDTGLEYANSYNPAKKENSPHLD 120  
QY 121 EYSIIQSMGYRRARARLLOSEPENPSLOETSLVOLSNLGTVRTLRKRIOPQKTSYVI 180  
DB 121 EYSIIQSMGYRRARARLLOSEPENPSLOETSLVOLSNLGTVRTLRKRIOPQKTSYVI 180  
QY 181 ELGSSSEEDTVNKAATYCSVGDOELQITPQGTREDEISLDSAKKACEFSEDTVTNHHQ 240  
DB 181 ELGSSSEEDTVNKAATYCSVGDOELQITPQGTREDEISLDSAKKACEFSEDTVTNHHQ 240  
QY 241 PSNNDLNTEKRAAERHPEKYGSSSVSNLHVEPCGNTNTHASSLOHNSSLLTKDMNVE 300  
DB 241 PSNNDLNTEKRAAERHPEKYGSSSVSNLHVEPCGNTNTHASSLOHNSSLLTKDMNVE 300  
QY 301 KAEFCNKSKOPGLARSQHNRMAGSKETCNDRTPTSEKRVLDNADPLCERKENKOKLPC 360  
DB 301 KAEFCNKSKOPGLARSQHNRMAGSKETCNDRTPTSEKRVLDNADPLCERKENKOKLPC 360  
QY 361 SENPDTEDEVPITLNNSTQKYNEMFSRDELLGSDSHDGESESAKAVADYDVTNEVD 420  
DB 361 SENPDTEDEVPITLNNSTQKYNEMFSRDELLGSDSHDGESESAKAVADYDVTNEVD 420  
QY 421 EYSSGSEKIDLLASDPHEALICKSERVHSKSVESNIEDKIFGKTYKKASLPVLHVTEH 480  
DB 421 EYSSGSEKIDLLASDPHEALICKSERVHSKSVESNIEDKIFGKTYKKASLPVLHVTEH 480  
QY 481 LIIGAFTVEPOIIOERPLTNLKKRRRPTSGLHPEDFIKKADLAVOKTPEMINOQTNOTE 540  
DB 481 LIIGAFTVEPOIIOERPLTNLKKRRRPTSGLHPEDFIKKADLAVOKTPEMINOQTNOTE 540  
QY 541 ONGOVMNTNSGHEKKTGDSIQNEKNPPIESLEKESAFKTKAEPISISSIMLELENI 600  
DB 541 ONGOVMNTNSGHEKKTGDSIQNEKNPPIESLEKESAFKTKAEPISISSIMLELENI 600  
QY 601 HNSKAPKNRLRRKSTRIHLELVSNLSPPNCTELQIDSCSSEETKKKKYNNOMPV 660  
DB 601 HNSKAPKNRLRRKSTRIHLELVSNLSPPNCTELQIDSCSSEETKKKKYNNOMPV 660  
QY 661 RHRNRLOMEGKEPATGAKKSNKPNQOTSKRHDSTFPPELKLTNAPGSETKCSNTSELKE 720  
DB 661 RHRNRLOMEGKEPATGAKKSNKPNQOTSKRHDSTFPPELKLTNAPGSETKCSNTSELKE 720  
QY 721 FVNPSLPREEKEKELEYKVSNNADPKDMLSGERVLOTERSVSSSISLVPDGYDGYO 780  
DB 721 FVNPSLPREEKEKELEYKVSNNADPKDMLSGERVLOTERSVSSSISLVPDGYDGYO 780  
QY 781 ESIISLLEVSTLGAKATEPNKCVSOCAAFENPKGLIHGCKDNRNDEGKRYPLGHEVNS 840  
DB 781 ESIISLLEVSTLGAKATEPNKCVSOCAAFENPKGLIHGCKDNRNDEGKRYPLGHEVNS 840  
QY 841 RETSIEMESELDAIYLONTFEVSKROSFALESNGNAEEGATESAHSGSLKOKSPKYT 900  
DB 841 RETSIEMESELDAIYLONTFEVSKROSFALESNGNAEEGATESAHSGSLKOKSPKYT 900  
QY 901 FECEOKKEENOGKNESNIKPVQTVNITAGFPVVGQKDKPDVNAKCSIKGGSRCFLSSQFRG 960

DB 901 FECEOKKEENOGKNESNIKPVQTVNITAGFPVVGQKDKPDVNAKCSIKGGSRCFLSSQFRG 960  
QY 961 NETGLITPKKHGLOVPYRIPPLPIKSPVKTKCKKNLLEEFHSHSPREKMNENIP 1020  
DB 961 NETGLITPKKHGLOVPYRIPPLPIKSPVKTKCKKNLLEEFHSHSPREKMNENIP 1020  
QY 1021 STVSTISRRNNIRENVFKEASSNINEVGSSTNEVSSINETIGSSDENIOAELGRNRGPKL 1080  
DB 1021 STVSTISRRNNIRENVFKEASSNINEVGSSTNEVSSINETIGSSDENIOAELGRNRGPKL 1080  
QY 1081 NAMLRLGLVLOPEYKOSLPGSNCKHPEIKQEEVVOFTVNTDESPLYISNLEQPMGSS 1140  
DB 1081 NAMLRLGLVLOPEYKOSLPGSNCKHPEIKQEEVVOFTVNTDESPLYISNLEQPMGSS 1140  
QY 1141 HASQVSETPDDLDDGELIKEDTSAENDIKRSSAVFESKSVOKGSELSRSPFTHTHQAQ 1200  
DB 1141 HASQVSETPDDLDDGELIKEDTSAENDIKRSSAVFESKSVOKGSELSRSPFTHTHQAQ 1200  
QY 1201 GYRGAKKLESSEENLSSDEELPCFOHLLFGKVNINISQSTRHSTVATECLSKNTEBNL 1260  
DB 1201 GYRGAKKLESSEENLSSDEELPCFOHLLFGKVNINISQSTRHSTVATECLSKNTEBNL 1260  
QY 1261 LSLKNSLNDCSNOVILIAKASQHHLSEETKCSASLFSQCCSELBDLTANTVTOBPFLIGS 1320  
DB 1261 LSLKNSLNDCSNOVILIAKASQHHLSEETKCSASLFSQCCSELBDLTANTVTOBPFLIGS 1320  
QY 1321 SKOMRHQESQOGVGLSDKELVSDDEERGTLGLENNOBQSDMSNLGEAASGESSTVSSE 1380  
DB 1321 SKOMRHQESQOGVGLSDKELVSDDEERGTLGLENNOBQSDMSNLGEAASGESSTVSSE 1380  
QY 1381 DCSGSSQSDILITTOQRTMOHNLIKLOEMAELAVLEQHSQSPNSYPSIISSDSALE 1440  
DB 1381 DCSGSSQSDILITTOQRTMOHNLIKLOEMAELAVLEQHSQSPNSYPSIISSDSALE 1440  
QY 1441 DLRNPEQSTSEKAVILTSQKSEFPISQNEGLSADKFEVSADSSTSKKKEGVESSPSK 1500  
DB 1441 DLRNPEQSTSEKAVILTSQKSEFPISQNEGLSADKFEVSADSSTSKKKEGVESSPSK 1500  
QY 1501 CPSLDRMYMHSCSGSLQNRNYPPOEBELIKVYDVEEQOLEESGPHDLTETSYLPRODLEG 1560  
DB 1501 CPSLDRMYMHSCSGSLQNRNYPPOEBELIKVYDVEEQOLEESGPHDLTETSYLPRODLEG 1560  
QY 1561 TPYLESGISLFSDDPESDPSEDRAPEARSVGNIPSTSAKVYPOLKVAESAQSPAANTT 1620  
DB 1561 TPYLESGISLFSDDPESDPSEDRAPEARSVGNIPSTSAKVYPOLKVAESAQSPAANTT 1620  
QY 1621 DTAGYNAMESVSREKPELTASTERYNKRMSWVSGLTPPEERMLVYKFAARHHITLNL 1680  
DB 1621 DTAGYNAMESVSREKPELTASTERYNKRMSWVSGLTPPEERMLVYKFAARHHITLNL 1680  
QY 1681 TEETTHVVKTDAAEFYCERETLKYFLGIAGGKVVSYFWVTOSIKERKMLNHEDEVRGDY 1740  
DB 1681 TEETTHVVKTDAAEFYCERETLKYFLGIAGGKVVSYFWVTOSIKERKMLNHEDEVRGDY 1740  
QY 1741 VNGRNHQGKRARESQDRKIFRGLEICCGPPTNPTDQLEMMVQLCASVVKELSSFTL 1800  
DB 1741 VNGRNHQGKRARESQDRKIFRGLEICCGPPTNPTDQLEMMVQLCASVVKELSSFTL 1800  
QY 1801 GTGVPRIVVOPDANTEDNGFHAIGQCEAPVYTHENLDSVALYQCOELDTYLIPOIPH 1860  
DB 1801 GTGVPRIVVOPDANTEDNGFHAIGQCEAPVYTHENLDSVALYQCOELDTYLIPOIPH 1860  
QY 1861 SHY 1863  
DB 1861 SHY 1863

RESULT 5  
US-09-982-828-2  
Sequence 2, Application us/09982828  
Publication No. US20030022184A1  
GENERAL INFORMATION:

APPLICANT: Murphy, Patricia D.  
 Allen, Antonette C.  
 Alvares, Christopher P.  
 Critz, Brenda S.  
 Olson, Sheri J.  
 Thurber, Denise  
 Zeng, Bin  
 TITLE OF INVENTION: Coding Sequences of the Human  
 BRCA1 Gene  
 NUMBER OF SEQUENCES: 72  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Morgan Lewis & Bockius LLP  
 STREET: 1111 Pennsylvania Avenue N.W.  
 CITY: Washington  
 STATE: DC  
 COUNTRY: USA  
 ZIP: 20004  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/982,828  
 FILING DATE: 22-Oct-2001  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 09/074,453  
 FILING DATE: 1998-05-06  
 APPLICATION NUMBER: US 08/798,691  
 FILING DATE: 1997-02-12  
 APPLICATION NUMBER: US 08/598,591  
 FILING DATE: 1996-02-12  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Michael S. Tuscan  
 REGISTRATION NUMBER: 43,210  
 REFERENCE/DOCKET NUMBER: 44921-5053-01-US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-739-3000  
 TELEFAX: 202-739-3001  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1863 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: not relevant  
 TOPOLOGY: not relevant  
 MOLECULE TYPE: protein  
 ORIGINAL SOURCE:  
 ORGANISM: Homo sapiens  
 STRAIN: BRCA1 (cm11)  
 POSITION IN GENOME:  
 CHROMOSOME/SEGMENT: 17  
 MAP POSITION: 17q21  
 SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
 US-09-982-828-2  
 Query Match 99.9%; Score 9635; DB 9; Length 1863;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 1860; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 181 ELGSDSSEDTVKNATYCSVGDELLQITPGCTRDEISLDSAKKAACESETDVTNTEHHQ 240  
 Qy 241 PSNNDLNTTEKRAERHPEKRYOGSSVSNLHJEPGNTNHAASLHENSILLTDNRNVE 300  
 Db 241 PSNNDLNTTEKRAERHPEKRYOGSSVSNLHJEPGNTNHAASLHENSILLTDNRNVE 300  
 Qy 301 KAECNKSQPGRLARSOHNRWAGSEKTCNDRRTSTEEKVLDNADPLCEKREWNKOKLPC 360  
 Db 301 KAECNKSQPGRLARSOHNRWAGSEKTCNDRRTSTEEKVLDNADPLCEKREWNKOKLPC 360  
 Qy 361 SENPRTDVEDVWITLNSIQKVNEMFSRSDLLCSDSHDESSNAKAVDVLNEVD 420  
 Db 361 SENPRTDVEDVWITLNSIQKVNEMFSRSDLLCSDSHDESSNAKAVDVLNEVD 420  
 Qy 421 EYSGSSEKIDLLASDPHEALICKSERVHKSVEENIEDEKIGKTYRKKASLPNLSHTEN 480  
 Db 421 EYSGSSEKIDLLASDPHEALICKSERVHKSVEENIEDEKIGKTYRKKASLPNLSHTEN 480  
 Qy 481 LIIGAFYTEPQIIQERPLTNKLRKRRPTSGLHPEDEFIKRADLAVOKTPEMINOGTNOTE 540  
 Db 481 LIIGAFYTEPQIIQERPLTNKLRKRRPTSGLHPEDEFIKRADLAVOKTPEMINOGTNOTE 540  
 Qy 541 QNGQVMNITNSGHNKTKGDSIQNEKNPNPIESLEKESAFYTKAEPITSSISNMELEINI 600  
 Db 541 QNGQVMNITNSGHNKTKGDSIQNEKNPNPIESLEKESAFYTKAEPITSSISNMELEINI 600  
 Qy 601 HNSKAPKKNRLRRKSSSTRHIALLVVSRNLSPPNCTELODSCSSSEIKKKRYNQMPV 660  
 Db 601 HNSKAPKKNRLRRKSSSTRHIALLVVSRNLSPPNCTELODSCSSSEIKKKRYNQMPV 660  
 Qy 661 RHSRNLQIMEGKEPATGAKKSNKPNEDTSKRHSDPTPELKLITNAPGFTCSNTSELEKE 720  
 Db 661 RHSRNLQIMEGKEPATGAKKSNKPNEDTSKRHSDPTPELKLITNAPGFTCSNTSELEKE 720  
 Qy 721 FVNPSPREEKEKLETVKVSNNAEDEPKDMLSGERVLOTFRSVSSSISLVPGTDTGTQ 780  
 Db 721 FVNPSPREEKEKLETVKVSNNAEDEPKDMLSGERVLOTFRSVSSSISLVPGTDTGTQ 780  
 Qy 781 ESISLLEVTSLGKAKTEPNKCVSOCAAFENKGLIHGSKNRRNDTEGFKPLDHEVHVS 840  
 Db 781 ESISLLEVTSLGKAKTEPNKCVSOCAAFENKGLIHGSKNRRNDTEGFKPLDHEVHVS 840  
 Qy 841 RETSIEMESELDQOYLQNTFEKVSROSFALEFSPNGNAEECAFTSAHSGSLKROSPKVT 900  
 Db 841 RETSIEMESELDQOYLQNTFEKVSROSFALEFSPNGNAEECAFTSAHSGSLKROSPKVT 900  
 Qy 901 FECBQKEENQKNESNIRPVQTVNITAGFPVVGKDKPVDNAKCSIKGSRFCLSSQFRG 960  
 Db 901 FECBQKEENQKNESNIRPVQTVNITAGFPVVGKDKPVDNAKCSIKGSRFCLSSQFRG 960  
 Qy 961 NETGLITPNKHGLQNTYRIRPLEPFIKSPYTKKKNULLENFEHSMSPREKNGENIRP 1020  
 Db 961 NETGLITPNKHGLQNTYRIRPLEPFIKSPYTKKKNULLENFEHSMSPREKNGENIRP 1020  
 Qy 1021 STVSTISRNRIRENVFKEASSNINEGSSSTNEVGSSSINETGSSDENIOAELGNNRQPKL 1080  
 Db 1021 STVSTISRNRIRENVFKEASSNINEGSSSSTNEVGSSSINETGSSDENIOAELGNNRQPKL 1080  
 Qy 1081 NAMRLGVLQPEVYKOSLPGSNCNHPETIKQOEVEVQVNTDPSYLLISDNLEQPMGSS 1140  
 Db 1081 NAMRLGVLQPEVYKOSLPGSNCNHPETIKQOEVEVQVNTDPSYLLISDNLEQPMGSS 1140  
 Qy 1141 HASOVCSSETDDLLDDEIEKEDTSFAANDIKESAVSKSVQKELSRSPFHTHLLAQ 1200  
 Db 1141 HASOVCSSETDDLLDDEIEKEDTSFAANDIKESAVSKSVQKELSRSPFHTHLLAQ 1200  
 Qy 1201 GYRGAKKLSSSEENLSSDEDELPCFOHLLFGKVNNPSPQSTRSYATFELSKNTEENL 1260  
 Db 1201 GYRGAKKLSSSEENLSSDEDELPCFOHLLFGKVNNPSPQSTRSYATFELSKNTEENL 1260  
 Qy 1261 LSLKNSINDCSNOVITLAKASQEHHLSEETKCSASLFFSSQCSLEEDLTANTNTQDPFLIGS 1320  
 Db 1261 LSLKNSINDCSNOVITLAKASQEHHLSEETKCSASLFFSSQCSLEEDLTANTNTQDPFLIGS 1320

```

Db      1261 LSLKNSLNDSCNOVILAKASQEHHLSEETKCSASLFSSQCELEDLTANTNODPELLIGS 1320
QY      1321 SKOMRHQSESGQVGLSDKELVSDDEBERGTGLENNNOEQQSMDSNLGEAASCESETSVSE 1380
Db      1321 SKOMRHQSESGQVGLSDKELVSDDEBERGTGLENNNOEQQSMDSNLGEAASCESETSVSE 1380
QY      1381 DCSGSSQSQSDILTTQORDTMQHNLIKLOQMALEVLLEQHGQSPNSVPSIISDSALE 1440
Db      1381 DCSGSSQSQSDILTTQORDTMQHNLIKLOQMALEVLLEQHGQSPNSVPSIISDSALE 1440
QY      1441 DLRNPEQSTSEKAVLTSQKSESEPISONPEGLSADAFEVSADSTSKNKPGEVRSRSPK 1500
Db      1441 DLRNPEQSTSEKAVLTSQKSESEPISONPEGLSADAFEVSADSTSKNKPGEVRSRSPK 1500
QY      1501 CPSLDRMYHSCSGSLQNRNYPQSEELIKVVDVEEQLEESGPHDLTETSYPRODLEG 1560
Db      1501 CPSLDRMYHSCSGSLQNRNYPQSEELIKVVDVEEQLEESGPHDLTETSYPRODLEG 1560
QY      1561 TPYLESGISLFSDDPESDESEDRAPESARVGNIPSSTSALKVPOLKVAESAQSPAAHTT 1620
Db      1561 TPYLESGISLFSDDPESDESEDRAPESARVGNIPSSTSALKVPOLKVAESAQSPAAHTT 1620
QY      1621 DTAGYNAMESVSREKPELTASTERYNKRMVSVGLTPEEFMLVYKFAKHHITLTNLI 1680
Db      1621 DTAGYNAMESVSREKPELTASTERYNKRMVSVGLTPEEFMLVYKFAKHHITLTNLI 1680
QY      1681 TEETHVAVKTADEFECERTLKFLGSIAGKVVVSYFWVTQSIKERKMLNEHDFEVRGVD 1740
Db      1681 TEETHVAVKTADEFECERTLKFLGSIAGKVVVSYFWVTQSIKERKMLNEHDFEVRGVD 1740
QY      1741 VNGRNHQGKRRARESDRKIFRGLEICYGPTNMPPTDOLWVVOGASVVELSSFTL 1800
Db      1741 VNGRNHQGKRRARESDRKIFRGLEICYGPTNMPPTDOLWVVOGASVVELSSFTL 1800
QY      1801 GTGVPIVYVODDAMTEDNGFHAIGMCEAPVYTRREMLVDSVALYOCQELDTLLIPIQIP 1860
Db      1801 GTGVPIVYVODDAMTEDNGFHAIGMCEAPVYTRREMLVDSVALYOCQELDTLLIPIQIP 1860
QY      1861 SHY 1863
Db      1861 SHY 1863

```

RESULT 6  
 US-09-982-828-4  
 : Sequence 4, Application US/09982828  
 : Publication No. US20030022184A1  
 : GENERAL INFORMATION:

APPLICANT: Murphy, Patricia D.  
 Allen, Antoinette C.  
 Alvares, Christopher P.  
 Critz, Brenda S.  
 Olson, Sheri J.  
 Thurber, Denise  
 Zeng, Bin

TITLE OF INVENTION: Coding Sequences of the Human  
 BRCA1 Gene

NUMBER OF SEQUENCES: 72  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Morgan Lewis & Bockius LLP  
 STREET: 1111 Pennsylvania Avenue N. W.  
 CITY: Washington  
 STATE: DC

COUNTRY: USA  
 ZIP: 20004  
 COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/982,828  
 FILING DATE: 22-Oct-2001

```

: CLASSIFICATION: <Unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 09/074,453
: FILING DATE: 1998-05-06
: APPLICATION NUMBER: US 08/798,691
: FILING DATE: 1997-02-12
: APPLICATION NUMBER: US 08/598,591
: FILING DATE: 1996-02-12
: ATTORNEY/AGENT INFORMATION:
: NAME: Michael S. Tuscan
: REGISTRATION NUMBER: 43,210
: REFERENCE/DOCKET NUMBER: 44921-5053-01-US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-739-3001
: TELEFAX: 202-739-3000
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1863 amino acids
: TYPE: amino acid
: STRANDEDNESS: not relevant
: TOPOLOGY: not relevant
: MOLECULE TYPE: protein
: ORIGINAL SOURCE:
: ORGANISM: Homo sapiens
: STRAIN: BRCA1 (om12)
: POSITION IN GENOME:
: CHROMOSOME/SEGMENT: 17
: MAP POSITION: 17q21
: SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-982-828-4

Query Match      99.9%; Score 9635; DB 9; Length 1863;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1860; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 MDSALRVEVQVNIAMKILIECPICLEIKPEVSTKCDHIFCFMKLLLNQKKGSG 60
Db      1 MDSALRVEVQVNIAMKILIECPICLEIKPEVSTKCDHIFCFMKLLLNQKKGSG 60
QY      61 CPCKNDITKRSLOESTRFSQLEVEELKLTICAFOLDTGEVANSYVFAKKNSEHKLKD 120
Db      61 CPCKNDITKRSLOESTRFSQLEVEELKLTICAFOLDTGEVANSYVFAKKNSEHKLKD 120
QY      121 EVSIIOGMYRNRAKRLLOSPEPNPSLOETSLSVOLSNLGTVRTLRTKQRIQPKTSVYI 180
Db      121 EVSIIOGMYRNRAKRLLOSPEPNPSLOETSLSVOLSNLGTVRTLRTKQRIQPKTSVYI 180
QY      181 ELGSDSEDTVKATYCSVGDELQITPQGTREDEISLSAKKACESESTDVNTTEHQ 240
Db      181 ELGSDSEDTVKATYCSVGDELQITPQGTREDEISLSAKKACESESTDVNTTEHQ 240
QY      241 PSNNDLNTTEKRAAERHPEKYOGSSVSNLHVEPGCTNHASLQHNSSLTLTDRMVE 300
Db      241 PSNNDLNTTEKRAAERHPEKYOGSSVSNLHVEPGCTNHASLQHNSSLTLTDRMVE 300
QY      301 KAEFCNKSQPGIANSQHNRMAGSKETCNDRTPTSTERRKVDLNDPLCERKEMKOKLPC 360
Db      301 KAEFCNKSQPGIANSQHNRMAGSKETCNDRTPTSTERRKVDLNDPLCERKEMKOKLPC 360
QY      361 SENPRDTEVPWITLNSIQKVENWFSRSDDELGSDSDHGESESNKAVADYLDVLEVD 420
Db      361 SENPRDTEVPWITLNSIQKVENWFSRSDDELGSDSDHGESESNKAVADYLDVLEVD 420
QY      421 EYSGSSEKIDLASDPHEALICKSRVYVSKSVESNIEDKIFGKYRKKASLPNLSHTEN 480
Db      421 EYSGSSEKIDLASDPHEALICKSRVYVSKSVESNIEDKIFGKYRKKASLPNLSHTEN 480
QY      481 LITGAFTVEPQITIOERPLTNKLRRKRPTSGLHPEDFTKADLAVOKTPEMINGNOTE 540
Db      481 LITGAFTVEPQITIOERPLTNKLRRKRPTSGLHPEDFTKADLAVOKTPEMINGNOTE 540
QY      541 ONGOVANITNSGHEKTKDSTIQNFKNNPILSEKESAFYTKAPLISSISNNELELNI 600
Db      541 ONGOVANITNSGHEKTKDSTIQNFKNNPILSEKESAFYTKAPLISSISNNELELNI 600

```

Db 541 QNCGVANNITNSGHNKTKGDSIONENKPNPDESLEKESAFKRAEPISSISIMMELNLI 600  
 QY HNSKAPKRNLRKRSSTRHIALELVYSRLSPNCLEIDIDSCSSSEELKKKKYNNMPV 660  
 Db HNSKAPKRNLRKRSSTRHIALELVYSRLSPNCLEIDIDSCSSSEELKKKKYNNMPV 660  
 QY 661 RHRNRNQLMEKBPATGAKKSNKPNQOTSKRHSDFPELKLTNAPGSPFKCNTSELKE 720  
 Db 661 RHRNRNQLMEKBPATGAKKSNKPNQOTSKRHSDFPELKLTNAPGSPFKCNTSELKE 720  
 QY 721 FVNPSPREKEKLELVYKVSNNADPKDMLSGERLQTERVSSISLVGTDTGTQ 780  
 Db 721 FVNPSPREKEKLELVYKVSNNADPKDMLSGERLQTERVSSISLVGTDTGTQ 780  
 QY 781 ESISLLEVTGLKAKTEPNKVSQCAAFENPKGLIHCCSKDNKNDTGFYPLGHEVNH 840  
 Db 781 ESISLLEVTGLKAKTEPNKVSQCAAFENPKGLIHCCSKDNKNDTGFYPLGHEVNH 840  
 QY 841 RETSIEMEESELDQYLONTFKVSKRSQSFALFSPNGAEECATFSAHSGSLKQSPKVT 900  
 Db 841 RETSIEMEESELDQYLONTFKVSKRSQSFALFSPNGAEECATFSAHSGSLKQSPKVT 900  
 QY 901 FECEOKEEENKRNESNKPQYVNTAGFPVYQKDKPVNMACSITKGSFCLSSQFRG 960  
 Db 901 FECEOKEEENKRNESNKPQYVNTAGFPVYQKDKPVNMACSITKGSFCLSSQFRG 960  
 QY 961 NETGLTPPNKHLQNPYRIPPLPIKSPYKCKKNLLEENFEHSMSPEREMGNINP 1020  
 Db 961 NETGLTPPNKHLQNPYRIPPLPIKSPYKCKKNLLEENFEHSMSPEREMGNINP 1020  
 QY 1021 STVSTISRRNIRENVFKEAASSNINEVGSSTNEVGSSINEIGSSDENIQAELRNRPKL 1080  
 Db 1021 STVSTISRRNIRENVFKEAASSNINEVGSSTNEVGSSINEIGSSDENIQAELRNRPKL 1080  
 QY 1081 NAMRLGVLOPEYKOSLPSSNCKHPKIKQOEYEVYQYVNTDPSPLISDNEQPMGSS 1140  
 Db 1081 NAMRLGVLOPEYKOSLPSSNCKHPKIKQOEYEVYQYVNTDPSPLISDNEQPMGSS 1140  
 QY 1141 HASQVSETPDDLLDGEIKEDTSFAENDIKESSAVFSKQKELSRSPPTHTHIAQ 1200  
 Db 1141 HASQVSETPDDLLDGEIKEDTSFAENDIKESSAVFSKQKELSRSPPTHTHIAQ 1200  
 QY 1201 GYRRGAKKLESSEENLSEDEELPCFOHLLFGKVNNTPSQSTRHSTVATBCLSKNTEENL 1260  
 Db 1201 GYRRGAKKLESSEENLSEDEELPCFOHLLFGKVNNTPSQSTRHSTVATBCLSKNTEENL 1260  
 QY 1261 LSTKNSLNCNSQVILAKASQEHLSBETKCSALFSSQCSLEDELTLANTNTODPELIGS 1320  
 Db 1261 LSTKNSLNCNSQVILAKASQEHLSBETKCSALFSSQCSLEDELTLANTNTODPELIGS 1320  
 QY 1321 SKQMRHOSQGVGLSDKELVSDDEERGTLLENNOEEOSSMDNGLGAAGCESETSYSE 1380  
 Db 1321 SKQMRHOSQGVGLSDKELVSDDEERGTLLENNOEEOSSMDNGLGAAGCESETSYSE 1380  
 QY 1381 DCSGLSSQSDILTTQORDTMQHNLIKLOOEMAELEAVLEOHSQSPNSYPSIISDSALE 1440  
 Db 1381 DCSGLSSQSDILTTQORDTMQHNLIKLOOEMAELEAVLEOHSQSPNSYPSIISDSALE 1440  
 QY 1441 DLNPNPOSTSEKAVLTSQKSESEPIQNPBGSLADFEVSADSTSKNKEPVERSSPSK 1500  
 Db 1441 DLNPNPOSTSEKAVLTSQKSESEPIQNPBGSLADFEVSADSTSKNKEPVERSSPSK 1500  
 QY 1501 CPESLDNRWYHSCGSLQNRNYPQOEBELIKVYDVECOOLESGPHULTETSYLPRDLEG 1560  
 Db 1501 CPESLDNRWYHSCGSLQNRNYPQOEBELIKVYDVECOOLESGPHULTETSYLPRDLEG 1560  
 QY 1561 TPYLESGISLFSDDPESDPESEDRAPESARVGNIPSSSTALKVPQLKVAESAQSPAHTT 1620  
 Db 1561 TPYLESGISLFSDDPESDPESEDRAPESARVGNIPSSSTALKVPQLKVAESAQSPAHTT 1620  
 QY 1621 DTAGYNAEMESVREKPELASTERVNRKMSMYVSGLTPEEFMLVYKFAKHHITTLNLI 1680  
 Db 1621 DTAGYNAEMESVREKPELASTERVNRKMSMYVSGLTPEEFMLVYKFAKHHITTLNLI 1680

QY 1681 TEETHVAMKDAEFVCEKTLKTEFLGIAGKMWVSFYFWTQSIKERKMLNEHDFEVBGV 1740  
 Db 1681 TEETHVAMKDAEFVCEKTLKTEFLGIAGKMWVSFYFWTQSIKERKMLNEHDFEVBGV 1740  
 QY 1741 VGNRNHOGPKRARESDQKIFRGLEICCYGPTNMPDQLEMMVYOLCGASVYELSSFTL 1800  
 Db 1741 VGNRNHOGPKRARESDQKIFRGLEICCYGPTNMPDQLEMMVYOLCGASVYELSSFTL 1800  
 QY 1801 GTGVHPIVYVQPAWTFEDNCFHAIIGOMCEAPVYTRENVLDSVALYOCQELDTYLIPIPH 1860  
 Db 1801 GTGVHPIVYVQPAWTFEDNCFHAIIGOMCEAPVYTRENVLDSVALYOCQELDTYLIPIPH 1860  
 QY 1861 SHY 1863  
 Db 1861 SHY 1863

RESULT 7  
 US-10-022-819-2  
 : Sequence 2, Application US/10022819  
 : Publication No. US20030027166A1  
 : GENERAL INFORMATION:  
 : APPLICANT: ALLEN, Antoinette C. P.  
 : OLSEN, Sheri J.  
 : LAWRENCE, Tammy  
 : ANGELLY, Tracy S.  
 : RABIN, Mark B.  
 : TITLE OF INVENTION: CODING SEQUENCE HAPLOTYPE OF THE HUMAN  
 : BCRA1 GENE  
 : NUMBER OF SEQUENCES: 67  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Morgan Lewis & Bockius LLP  
 : STREET: 1111 Pennsylvania Avenue  
 : CITY: Washington DC  
 : STATE: District of Columbia  
 : COUNTRY: USA  
 : ZIP: 20004  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: PatentIn Release #1.0, Version #1.30  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/10/022,819  
 : FILING DATE: 22-Apr-2002  
 : CLASSIFICATION: <Unknown>  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: 09/074,452  
 : FILING DATE: 1998-05-06  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: <Unknown>  
 : REGISTRATION NUMBER: <Unknown>  
 : REFERENCE/DOCKET NUMBER: 044921-5049-01-US  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: 202-739-3000  
 : TELEFAX: 202-739-3001  
 : INFORMATION FOR SEQ ID NO: 2:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 1863 amino acids  
 : TYPE: amino acid  
 : STRANDEDNESS: single  
 : TOPOLOGY: linear  
 : MOLECULE TYPE: protein  
 : HYPOTHETICAL: NO  
 : ANTI-SENSE: NO  
 : FRAGMENT TYPE: N-terminal  
 : ORIGINAL SOURCE:  
 : ORGANISM: HOMO SAPIENS  
 : STRAIN: BCRA1  
 : HAPLOTYPE: OM14  
 : POSITION IN GENOME:  
 : CHROMOSOME/SEGMENT: 17

MAP POSITION: 17q21  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-10-022-819-2

Query Match 99.8%; Score 9627; DB 9; Length 1863;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1859; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

QY 1 MDLSALRVEEONVIAAMOKIIECPICELIKPEVSTKCHIFCKKCMKLLNOKKQPSQ 60
DB 1 MDLSALRVEEONVIAAMOKIIECPICELIKPEVSTKCHIFCKKCMKLLNOKKQPSQ 60
QY 61 CBLCKNDITKRSLOESTRESQVLEELIKICAFQDLTGLEYANSYNFAKKENNSPEHLK 120
DB 61 CBLCKNDITKRSLOESTRESQVLEELIKICAFQDLTGLEYANSYNFAKKENNSPEHLK 120
QY 121 EYVSIISMGYRRNRAKLLQSEPNPSLOETSLSVQLSNLCTVPTLRKQIQOKTSVYI 180
DB 121 EYVSIISMGYRRNRAKLLQSEPNPSLOETSLSVQLSNLCTVPTLRKQIQOKTSVYI 180
QY 181 ELGSDSEEDTVNKATYCSVGDOELLQITPOGTREDEISLDSAKKACEFSETDVTNTEHHQ 240
DB 181 ELGSDSEEDTVNKATYCSVGDOELLQITPOGTREDEISLDSAKKACEFSETDVTNTEHHQ 240
QY 241 PSNNDLNTTEKRAERHPEKYQSSVNLHVEPCGNTFHASLQIHENSLLTKDRMNV 300
DB 241 PSNNDLNTTEKRAERHPEKYQSSVNLHVEPCGNTFHASLQIHENSLLTKDRMNV 300
QY 301 KAEFCCKSKOPGLARSOHNRNAGSKETCDNRRTPTTEKKYVDLADPLCEKKEVNNKOLPC 360
DB 301 KAEFCCKSKOPGLARSOHNRNAGSKETCDNRRTPTTEKKYVDLADPLCEKKEVNNKOLPC 360
QY 361 SENPRDEDPVITLLNLSIQKVEMFSRDELLGSDSDHGESESNAKVADVLDVLEND 420
DB 361 SENPRDEDPVITLLNLSIQKVEMFSRDELLGSDSDHGESESNAKVADVLDVLEND 420
QY 421 EYSGSSEKIDLLASDPHEALICKSERVSHSKVESNIEDKIFGTYRKASLPMLSHVTEN 480
DB 421 EYSGSSEKIDLLASDPHEALICKSERVSHSKVESNIEDKIFGTYRKASLPMLSHVTEN 480
QY 481 LITGAFVIEPQIIQERPLTNKLRKRRTSGLHPEDFIKKADLAVOKTPEMINOGNOTE 540
DB 481 LITGAFVIEPQIIQERPLTNKLRKRRTSGLHPEDFIKKADLAVOKTPEMINOGNOTE 540
QY 541 QNGQVUNITNSGHEKNTKGDSIONEKNPNPIESLSEKESAFKTAEPDISSISIMMETELNI 600
DB 541 QNGQVUNITNSGHEKNTKGDSIONEKNPNPIESLSEKESAFKTAEPDISSISIMMETELNI 600
QY 601 HNSKAPKKRRLRRKSTRIHALELVYSRNLSPPNCTELOIDSCSSSEIIRKKKYNQMPV 660
DB 601 HNSKAPKKRRLRRKSTRIHALELVYSRNLSPPNCTELOIDSCSSSEIIRKKKYNQMPV 660
QY 661 RRSRNLOLEMEKEPATGAKKSNKPNQTSKRHSDTFPELKLTNAGSFTKCSNTSELKE 720
DB 661 RRSRNLOLEMEKEPATGAKKSNKPNQTSKRHSDTFPELKLTNAGSFTKCSNTSELKE 720
QY 721 FVNPSLPREKEEKLLETYSVSNNAEDPKDMLSGEYVLOTERSVESSISLVBGTDTYGTQ 780
DB 721 FVNPSLPREKEEKLLETYSVSNNAEDPKDMLSGEYVLOTERSVESSISLVBGTDTYGTQ 780
QY 781 ESIISLLEVSTLCKATKTEPKKVCYSCAFAENPKGLIHGCSKDNNDTEGFRYPLGHEVNS 840
DB 781 ESIISLLEVSTLCKATKTEPKKVCYSCAFAENPKGLIHGCSKDNNDTEGFRYPLGHEVNS 840
QY 841 RETSLEMESELDAAOVLONTFKVSKROSFALFSPNPAEBCATFSAHSGSLKQSPKYT 900
DB 841 RETSLEMESELDAAOVLONTFKVSKROSFALFSPNPAEBCATFSAHSGSLKQSPKYT 900
QY 901 FECCEKKEENOGKNESNIKPVQTVNITAGFPVVGQKDPVDMNAKCSIKGSRFLCSQFPG 960
DB 901 FECCEKKEENOGKNESNIKPVQTVNITAGFPVVGQKDPVDMNAKCSIKGSRFLCSQFPG 960
QY 961 NETGLITPKKHGLLONPYRIPLPLFIKSFVTKCKKKNLLEENFEHSHMSPEREMGENIP 1020

```

```

DB 961 NETGLITPKKHGLLONPYRIPLPLFIKSFVTKCKKKNLLEENFEHSHMSPEREMGENIP 1020
QY 1021 STVSTISRRNIRRENVKKEASSNNINEVSGSTNEVGSINIEISSDENTQAEIGRRRGKL 1080
DB 1021 STVSTISRRNIRRENVKKEASSNNINEVSGSTNEVGSINIEISSDENTQAEIGRRRGKL 1080
QY 1081 NAMLRGLVLOPEYKQSLGCSNCKHPEIKKOEVEEVQVJVNDFSPYLLSDLEQPMKSS 1140
DB 1081 NAMLRGLVLOPEYKQSLGCSNCKHPEIKKOEVEEVQVJVNDFSPYLLSDLEQPMKSS 1140
QY 1141 HASQVSETPDDLLDDEIKEDTSAENDIKSSSAVFSKQVKGELSNRSPFTHHLAQ 1200
DB 1141 HASQVSETPDDLLDDEIKEDTSAENDIKSSSAVFSKQVKGELSNRSPFTHHLAQ 1200
QY 1201 GYRRGAKKLLESSSEENLSSDEBELPCFHLLFGKVNIPQOSTRHSTVATECLSKNTEENL 1260
DB 1201 GYRRGAKKLLESSSEENLSSDEBELPCFHLLFGKVNIPQOSTRHSTVATECLSKNTEENL 1260
QY 1261 LSLKNSLINDCSNOVILAKASQEHHLSEETKCSASLFSQCSLEDLTANTQDPFLIGS 1320
DB 1261 LSLKNSLINDCSNOVILAKASQEHHLSEETKCSASLFSQCSLEDLTANTQDPFLIGS 1320
QY 1321 SKQMRHQSSEQVGLSDKELVSDDEERGTLLENNOEBSMDSNLGEAAGCESETSVSE 1380
DB 1321 SKQMRHQSSEQVGLSDKELVSDDEERGTLLENNOEBSMDSNLGEAAGCESETSVSE 1380
QY 1381 DCSGLSSQSDITLTQOQDPMOHNLIKLOEMLELEVLQHOQSOPNSPSTIISSSALE 1440
DB 1381 DCSGLSSQSDITLTQOQDPMOHNLIKLOEMLELEVLQHOQSOPNSPSTIISSSALE 1440
QY 1441 DLRNPEOSTSEKAVILTSQKSESEYPISONPEGLSADKFEVSAOSSTSKNKEPVERSSPSK 1500
DB 1441 DLRNPEOSTSEKAVILTSQKSESEYPISONPEGLSADKFEVSAOSSTSKNKEPVERSSPSK 1500
QY 1501 CPSLDDRWYMHSCSGSLQNRNTPSOBELIKVVDEBQOLEESGPHDLTETSYLPQDLEG 1560
DB 1501 CPSLDDRWYMHSCSGSLQNRNTPSOBELIKVVDEBQOLEESGPHDLTETSYLPQDLEG 1560
QY 1561 TPYLESGISLFSDDPSPSEDPAPASAVGNIPSTSLKVPOLKVAESAGPAAHTT 1620
DB 1561 TPYLESGISLFSDDPSPSEDPAPASAVGNIPSTSLKVPOLKVAESAGPAAHTT 1620
QY 1621 DTAGYNAMEESYREKPELTASTERYNKRMSVNVSGLTPPEEMLYYKPARKHHTLTMLI 1680
DB 1621 DTAGYNAMEESYREKPELTASTERYNKRMSVNVSGLTPPEEMLYYKPARKHHTLTMLI 1680
QY 1681 TEETTHVVKTDAAEFVCERTLKYFLGIAGKVVVSYFWVTOSIKERKMLNEHDFEVRGDV 1740
DB 1681 TEETTHVVKTDAAEFVCERTLKYFLGIAGKVVVSYFWVTOSIKERKMLNEHDFEVRGDV 1740
QY 1741 VNGRNHQGPKRARESDDRKIFRGLEICCYGPTNPTQLEMMVOLGASVYKELSSFTL 1800
DB 1741 VNGRNHQGPKRARESDDRKIFRGLEICCYGPTNPTQLEMMVOLGASVYKELSSFTL 1800
QY 1801 GTGVHPDIVVQDPDWTEDNGFHAIGMCEAPVYTBEMVLDVALYQCCQELDTYLLPQIPH 1860
DB 1801 GTGVHPDIVVQDPDWTEDNGFHAIGMCEAPVYTBEMVLDVALYQCCQELDTYLLPQIPH 1860
QY 1861 SHY 1863
DB 1861 SHY 1863

```

RESULT 8  
US-09-815-242-12713  
; Sequence 12713; Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Karl L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel



QY 1622 -----TAGY-----NAM-----EESVSRKPELTASTERY 1646  
DB 2170 STNSGSMNSIRHFTSLTSLGSLMSDSDSVISASESMASMSDSMSDST 2229  
QY 1647 NKRMNMYVSGLPFEEMLVYKFAKH-----HITLNLITEET----- 1684  
DB 2230 SSSMSNSMSMSTSEENSI-----HPSMSMSTQTHFTSTSTSESIAPNTNESOSTL 2282  
QY 1685 --THVVMKTDAE 1694  
DB 2283 SATSVSKKHDAE 2294  
RESULT 9  
US-09-815-242-12996  
Sequence 12996, Application US/09815242  
Patent No. US20020061569A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
TITLE OF INVENTION: Identification of Essential Genes In  
FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 12996  
LENGTH: 6281  
TYPE: PRT  
ORGANISM: Staphylococcus aureus  
US-09-815-242-12996  
Query Match 3.5%; Score 341; DB 10; Length 6281;  
Best Local Similarity 16.9%; Pred. No. 3.3e-09;  
Matches 389; Conservative 347; Mismatches 808; Indels 762; Gaps 90;  
QY 7 RVEEYQVYINAMOKLCEPICLEIKPEVSTKCDHFCKMLKLLNKKGSQCELCN 66  
DB 2848 QVEAAMQVNAATKALNGLNQLNLEKAKQAHNTAID-----GSLHLLTNAK 2891  
QY 67 DITKSLQESTFSLQVEELLKIICAFOLDGLE-----YANSYNFAKKENNSP 115  
DB 2892 EALKQLVQGSTTVAAGQNEQ---ANNVDAMDKLRQSIADNATTKQONNTDASQNK 2948  
QY 116 EHLKDEVSIIISMGYRNRAKRLQSEPNPSIQETSLVSQLSNLTGR-TLRTKQRIQO 174  
DB 2949 DAYNAAYTTAAG-----IIDQTSPTLDPVINOAGQVSTTKNALNGNELEAA 2998  
QY 175 KTSYIELGSDSSSEDTYKATYCSYGD-----ELLQI-----TPQGT-RDEIS 217  
DB 2999 KOQASQSIG---SLDNLNNAKQOTVTDQINGAHYVDEANQIKQONQNLNTAMGLNKQATA 3055

QY 218 LOSAKKACEFSETD-----VTNTEH----- 238  
DB 3056 DKDATKATVNFETDADQAKQOAVNTAVTNAENIISKANGNATQAEVEQAIKQVNAKQAL 3115  
QY 239 -----HQP-----SNNDLNTTEKRAARHREKQOGSSVSNLHREPGCTNTHASL 283  
DB 3116 NGNANYQAHKDEATALINSSNDLNOAQDAL-----KOQVQATTV-----AGVNVKQTA 3166  
QY 284 OHENSSLLLTK-----DRMN-----VEKAFCNKRK-----OPGL 313  
DB 3167 QELNANMTQLOKGIADKEQTRKADGNFVNAADPRKQANAYQAAKAEALLSATPDVYVTPSE 3226  
QY 314 ARSQHNRWAGSKETCNDRRTPSTERR-VDLNADPL-----CERKEWNKQLPCSENR-D 366  
DB 3227 ITAALNKVTQAKNDLNGNNTNLATAKQNVQHAIDLQPNLNQAKORDEYSKQIQTATLVPRVN 3286  
QY 367 TEDVPWITLINSILOKYNENFERSDELGSDSDSHDESE-----SNA----- 407  
DB 3287 AIQQAATVTLNDAMTQLOKGIANKAQIKSENYHDADTDKQAYADNAVYKAEELLKQTTNP 3346  
QY 408 -----KVADVLVLNVEDEYSGSSER-----IDLASDPHEALICKSERVH 448  
DB 3347 TMDPNTIDQALTKVNDTQNALNGNOKLADAKODATTTGTLDDHNDQAKQLTTQVEQAP 3406  
QY 449 SKSVESNIEDKIFGKYRKASLPMI.SHVTEMLIIGAFVTEPQIIQERPLT----- 499  
DB 3407 DIATVNVN-----KQNAQNLNNAMTNL-----NNMLQDKTEFLNSINFDA 3447  
QY 500 NKIKRRKRPSTGILHEDFLTKAD-----LAVOKTPEMINQSTNQTEDNGVNM----- 547  
DB 3448 DQAKDAYTNAVASHAGILSKANGNASQTEQAMQVNEAKQALNGNDVQBAKDAK 3507  
QY 548 --ITNSGHENKTKGDSI-----QNEKNPNPIE----- 572  
DB 3508 QVITNANLNOAQKQALQOYDAQTVANVTIKQTDADLQAMTQLOKGIADBDQTKAN 3567  
QY 573 -----SLEKESAFK--TKAEPISSISNMEL-----ELN----- 599  
DB 3568 GNFVNAADTKQNAYNNAVAHAEOIISGPRNANVDQQAQALQOQNAKGDNLGNHNLQV 3627  
QY 600 -----IHNKARKKKNLRKKSSTRIHAEIVYSKNLSPPNCTEL----- 639  
DB 3628 ARDNANTALDQLPNLNQPKTALKDQVS---HA-ELTVGNVAIKQNDALNNAMGTLKQ 3682  
QY 640 QIDSCS-----SSEPIKKKYNN-----QMPVHRSNL 666  
DB 3683 QIQANSQVYQSVDFQADQDQKQOAYNNANAOQIANGIPTVLTPPTVYQAVTTMQAK 3742  
QY 667 QLMGKEPATGAKKS-----NKNPN-----EOTSKR----- 691  
DB 3743 DALNGDEKLAQAKQOELANLDTLRDQNPQRDALRNQINQAOALATVEQTRKQNNQVNTA 3802  
QY 692 -----HDSDFPELKLUTNAPGSTKCSNSELKEFVNSLPREK 731  
DB 3803 MSNKLQGIANKDKVAKASENYHDADQKQATYTNVSAQEGILN-----QTNPLNLPDEI 3857  
QY 732 EEKLETFVAKSNAEDPKDMLSGSERVLQTER-----SV 764  
DB 3858 TRLAQVYTDKNG-----LNGEAKLTERQNAKDAVSGMTHLNDQAKQALQOIQSP 3910  
QY 765 ESSSISLVPCTDYGTQESISLLEVSTLGKAT-----EPNKCVSQCAAFENPGLI 815  
DB 3911 ELATVNOVKQATATSLDQAMDQDSQAINDKAQTLADGNYLNDPDKQNNAYQAVAKABALL 3970  
QY 816 HGCSDKNRNDTEGEGFYPLGHEVYNSHRET---SIEMESELDA-QYLONTKRV--SKRQSF 869  
DB 3971 N--KQSGTNEVOAQVESITNEVNAKQALNDNRLANNAKQOAKQOALNTHLNDQAKOSF 4028  
QY 870 -----ALFSNPGNAEECAFFSAHSGSLKQOS---PKVFEDE-----QKRENO 910  
DB 4029 ESQITQAPLVTVDTYTTINQKQTLID-HAEELLRNSVADNQTTLASEDYHDATPAORONDYNO 4087



```

QY 911 GKNEINIKPVQTVNITAGFPVVGOKDKPVNDNAKCSIKGSRFCLISQFRGNETGLITPNK 970
D 4088 AVTAANNIINQTSPTMNPDPVANGATTQVNNKTVALDGDENLAAAOQAN----- 4138
QY 971 HGLLONPRIPLPLPKSEVYTKCKKNLEEN-----FEHSMSPEREMN----- 1016
D 4139 -----RLDQDHDHNNMAOKOOLQOSLITQSSDIAAVNGHKQTAESINTAMGNLINAID 4190
QY 1017 -ENIPSTVSTISNNINRENF-----KEASSNINENGSSN-----EVGSSINIEISS----- 1064
D 4191 HQAVEORGNITNDTDQATVNTAVNEAAMINKQTOGNANQTEVQALTKVQTTQALN 4250
QY 1065 -DENIOAELGRNRPKLMLRLGLVLOPEYK-----OSLPG 1100
D 4251 GDNHNLQV-AKTNTQALDALTSLNDPOKTLKDQVTAATLVAVHOIEQANLNLQAMHG 4309
QY 1101 -----SNCH-----PEIKQEEVYQ-----TV 1120
D 4310 LROSIODNATKANSKYINEDPE--QONVQAVOANNIINQOTATLNNAINQATV 4367
QY 1121 NTFESPYSLLDNQPMGSSHASOVSETP-----DILLDGE-----IKED----- 1162
D 4368 NT--TKAALGDVKLONDKDHAKQTVSOALHNNACKHMETDILDETTTAVKQDLTEA 4425
QY 1163 -----TSFANDIKESSAVF-----SKSVQKGE 1185
D 4426 QALDQMLDALQOSIADKATRASSAYVNAEPNKQSYDEAVNAESTIAGLNPITINKGN 4485
QY 1186 L-SRSPSPPTHTHHLAOCYRRGAKKLESSEENLSEDEELPCFOHLLFGKYNIPSOSTRH 1244
D 4486 VSSATQAVISSKNAIDGVERLADQKOTAGNSLNHLQDLTPPAOQALENOJINN----- 4537
QY 1245 STVATFCLSKNRENTL-----SLKNSLNCDSNOVILAKRSQGHLSSEFKCASLFSQ 1299
D 4538 ATRRGEVAOKLTBAQLNOMEALRNSIODQOTEGSK-----FINEDKPOKDAYQAA 4591
QY 1300 CSELEDLITANTN--TQDPELLIGSSQMRHOSQSVGLSDKELVSDDEERGTLSE----- 1353
D 4592 VQNAKDLINQNTPTLDKQVEQLQAVNQAKNLHG--DQKLADKQNAVTDLNQL 4649
QY 1354 NNOEBSMDSNLGEASGE-----SETSY 1378
D 4650 NNPORALQSQJINNAATRGVNAOKLAFAKALDOAMALRNSIODQOTEGSKFINEDKP 4709
QY 1379 SEDC--SGLSQSDILITQOQDRTMOHNLKLOEMAELEAVLEQHSQ--SNRYPTISD 1435
D 4710 QKDAYOAAVQNAKDLINQNTGNPTLDKSOVE-QLTQAVTTAKONLHGDKLARDQQAQVTT 4768
QY 1436 SSALEDLRNPEQSTSEKAVLTSQSSSEYPISONPEGLSADKFEVSADSSSTKKN----- 1489
D 4769 VNALPRLNHAQOQALDAINAAPTRE--VAQHVQAT--ELDHAMELTKKXVQOVNT 4822
QY 1490 ---EPGEVSSPCKCSLDD---RWYHSCGSQONRNPYSQOELIKVYDVEQOLEE 1541
D 4823 DKAPYVTEASTDKKAVDAQLOAAESTIDPTNGSNANKDAYDQ---VLTKEKENEL 4878
QY 1542 SCPHDLTETSYLPRODEGTPYLESISLFSDDPESDPSDRPESARVGNIPSSSALK 1601
D 4879 NGNERVAEKTKAQKQITIDLTILNAD-QIATAKONIDQATKLOPIALEVDQATOLNOSMD 4937
QY 1602 VPOLKAESASQSPAHAHT--DTAGYNAMESVSRKPEL--TASTERYNKRMSAVS-- 1655
D 4938 QLOQAAVNEHANVQOTVTOADSDKONAKYKQALADENVLKQANQOVDQALONLNNK 4997
QY 1656 -GITPEEFMLVYKFAKKHITLTNLI 1680
D 4998 QALNGDERVALAKTKNGKHIDOLNAL 5023

```

RESULT 10  
US-10-171-311-4  
; Sequence 4, Application US/10171311  
; Publication No. US20030087270A1

```

; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatz, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoerst, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; FILE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; PRIOR FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 3899
; TYPE: PR
; ORGANISM: Homo sapiens
; US-10-171-311-4

Query Match 3.5%; Score 333; DB 9; Length 3899;
Best Local Similarity 18.5%; Pred. No. 4.6e-09;
Matches 393; Conservative 337; Mismatches 775; Indels 622; Gaps 93;

QY 7 RVEVQNVINAMQKILECPICLIEKPVSTKCDHFKCKMILKLNQKGPSCPLCKN 66
D 1339 QVQELSLISSLO-----QOLKETQNYEAELHCLQKRLQAAVSESTVPSLPDVS 1389
QY 67 DITKRSLOESTRPSQVLEELTKITCAFDLPGLGVANSYFAKKENN-----SEPHL 118
D 1390 VITESDQRTMYGSCVKK-----NIDGTIEFSGEG-VKEETIVYKLEKQTOEOL 1440
QY 119 KDEVS-IIOSM-----GYRNRA--KRLQSEPNPSLOETSLSVOLSNL 159
D 1441 EEEVAVIYMSJAFNAQOTELSLSGKENTASSKQAHAAVCOQOEHFNEMKLSQDQIGF 1500
QY 160 GIVRT--LRTKQRIQPKTSVYIELG-----SDSSEDTVNAKATYCSVDQELLQ- 206
D 1501 QTFETVDVKKKEEFKPLSK-----ELGEGKEILLNSDPDHDIPESKCVLTISEMFSKD 1556
QY 207 ---ITPOGTDEL---SLDSAKKACEFSETD-----VTNTEHQPNSNDLMTTERRA 254
D 1557 KTFIVRQSHIDELSVSSMDASRQMLNEQLEDMROELVQOYEHQOAFELLRQAHWRM 1616
QY 255 ERHPEKYOGSSVNLHVEPCGTN---THASSLOHENSLLITFDRMVEKAFCNKSQOP 311
D 1617 ERQREDQE-----QLQELIKRLNRQLAQRSSIDNEN--LVSEKRYVLEBLE--ALKOL 1666
QY 312 GLARSQHNKAGSKETPCNDRPTSTEEKVDLADPLCEKEMKOKLPCSENPREDVDV 371
D 1667 SL-----AGREKLCCELRNLSSTQTO---NGNE--NQEVEEVEQTFKEKELRKPEDVP 1713
QY 372 WITLNS--SIOKVN-----EWSRSDELIGSDSHGSESNAKV- 409
D 1714 PELLSERTALQANRBLKILLEVYKTTAAVEETIGR--HVLGIID-RSSKSQSSASLI 1770
QY 410 ---ADVLDVNE-----VDYSGS----- 425
D 1771 WRSEAEASVKSCHHEHTRTYDESIPSYSGDPRNDINMKSVPTEGTELSQRLVRSQF 1830
QY 426 -----SEKIDLLAS-----DPH 437
D 1831 AGTEIDPENELMLNITSSRLQAAVERKLLEAIFETSSQLHAKVQOTELMRESFROKQEA 1890

```

```

QY 438 EALICKS---ERVSHS-----VESNIEDK-1FGTKYRKASLPMLSHV 477
Db 1891 ESLKOEELREHEESRAROEALVELSKAEQVIDYADEKTLFEHQIEKTDI--IDRL 1948
QY 478 TENLIGA-----FVTEPQIIIOERPLTKLKRKRPISG-----LHEDFIKKADLAV 525
Db 1949 EOEELCASRRLOELEAEQOOIOEERELLSROK EAMAEAGPVEQOOLLOETEKLMKRELEV 2008
QY 526 OKTPMINOGTNOTONGVNMNITNSGHNKTKGDSIQNEKNPNPIESLEKESAFITKAE 585
Db 2009 OQQAQKAVDDLOK-----QVKALEIDVEQVSRFIELBOEKNTLMDLNOQNALEKOLE 2063
QY 586 PI-----SSSISMELELNI-----HNSAPRK--NRLRKSS 616
Db 2064 KMRKELDRAIDREHERDVFQOEIOKLEBOQLKVPRFOPISEHQTREREQALNHLKED 2123
QY 617 TRHIALELVASRNLSPPKCTELOIDSSSESEIKKKYKNOMPVRHSRLNLOMEGKE--P 674
Db 2124 -----KCELLLSKE-----OLORDIOERNEETIELEFR--VRELEQALLVEDKRRHG 2169
QY 675 ATGAKSKNPNBOTSRRHSDPTFPELTKLTNAPG-----SFTK 711
Db 2170 AVEAKPELSLEVOQLAERDAIDRKEKEITNLEQOLEFRELLENKNEVEVOQLHMOLEIQ 2229
QY 712 CSNSELKEFVNPSLPREEKEKLE-TVKVSNNADPKDMLSGE--RVLOTERSVESSS 768
Db 2230 KESTRILOLEOENKLFKDMKELGLAIKESDAMSTODQHVLFGKPAQIIQ--EKEVEIQ 2288
QY 769 ISLVGCTYGTQESISLE-----VSTLGAKEPKNKCVS-----QCAFENKGLIHGC 818
Db 2289 LN-----EQYTKLOOOLKITTDKNVIEEKNELINDLETOIECLMSDOB-----C 2332
QY 819 SKDNND-----TEGFKYPLGHEVNHSHRETSIEMESELEDAQYLOMTFKVSKROSPA 870
Db 2333 VKRNEEELIEQNEVIEKLOQEL---ANIGOKTSMNAHSLSEADSLKHQLODVIAEKLA 2389
QY 871 LFSNGNAEBECATSAHSGSLKOSPKVTFPECQ-----KEENOKRNESNIKPVOT 922
Db 2390 LEOQVETANEE-MTF-----MKNVLKETNFKMNQTOELFSLKRESEVEKIOSIPENS 2442
QY 923 VNITAGFVVGOKDR-----VDNAKCSIKGSRFCLSOFRGNFTGLITPKHGLLN 976
Db 2443 VNVALDHL---SKDRPELEVLTEDLKSLENOTYF---KSEENKGSITINLETRLO- 2495
QY 977 PYRIPLPEIFKSVTKCKKNLLE-----ENFEHSHSPEREMENIPSTVSTIS 1027
Db 2496 ---LESTVASAKDELTOCKQKQIKDMQEQOFETMLQKIVNLQKIVEKVAALVSOQL 2552
QY 1028 RNNIRE-----NFKREASSNI-----NEVGSTNEVGSSINEIGSSD----- 1065
Db 2553 LEAVOEYAKFCODNOTISSEPERTNIOMLQREDELGSDIALTLRLISELSQVAVEMHT 2612
QY 1066 -----ENIOAELG-RNRGPKLNLMLRGVLOPEVYKOSLPGSKCKHPEIKOEVEEVOVY 1120
Db 2613 SLIEKEVEYLAERKAVLEKEKKLEL-----OKLEGNEKKOREKKKSSPOVEVL 2664
QY 1121 NTDESPLYLIDNLE---QPMGSSHASOVCSETPDDLDDGEIKEDTSEAFENDIKESSAV 1176
Db 2665 KT--TTELFHSNEESGFNELEALRAESVATKA-----ELASVKEKAERKQOEL 2711
QY 1177 FSKSYQKELSRSPFTHTHLAOGYRGCAKLESSEENLSEDEBLPCFOHLLFGKVVN 1236
Db 2712 LVKETNMTSLQKDLQ--VRDHLAEA--KEKLSILEKEDETEVEQSKACFEELPIKLSKS 2769
QY 1237 IPSQSTRSTVATELSKNTENLILKNS-----LNCSSNOVILAKASQ----- 1281
Db 2770 IASQ-----TQGTLKISSNQPOLLVKNAGIQIINQSSCSSEVEYELISQTEKIEKMO 2824
QY 1282 -----EHLSE-----ETKCSALFSSOCSELEDTANTMTODPLIGSS--KO 1323
Db 2825 ELHAAELIDMESRHSISETETLKREHYVAVQLLKECGTLKANVIOCLRSK---SSSIFE 2880
QY 1324 MRHOSQSGVGLSDKELYSDDEERGTLGEENNQOEQSDMSNLGEAASGCESET---SVSE 1380

```

```

Db 2881 LAHSDAVQ-----TREICSSDS--CSDWOCGITYLTHSGCFIASEGRCSEESATDPSFK 2933
QY 1381 DCSGLSSOSDILLTQOORDTMQNLNLIKQOEMAEFAVLEOHGSP--SNSYSIISDSSAL 1439
Db 2934 KIKGL-----LRAVNBGMQ--VLSLTPSYSDGEDHSIQOVSSEPMLEEKRAVINTLSSL 2986
QY 1440 EDLRNEQOSTSEKAVLTSQSSSE-----YPISONPEGISADKFEVSADSS 1484
Db 2987 KDLITKMQQRAEAVYDSSQSHESFSDMRGELLALQOVLFEERSVLLAAFRTELTLALGT 3046
QY 1485 T-----SKNPEGYERSSPSKCSLDRMTMHCSSGLONRNPSPDELI----- 1529
Db 3047 TPAVGLNLEORIEOGVEYQAAECLOKADR-----RSLSEIOALHAQMN 3094
QY 1530 -KVVDEQOLESFGHDFTSTYLPRODLEGTPLYESGISLFPDPSDPSDEAPRSA 1588
Db 3095 GKRIITKREOESKPSOELEFINIOQO-----SOMLEMOVELS-----MKDRATELO 3143
QY 1589 RVGNIPSSSTALKVPOLKVAESQSPAARTDTAGYNAMEB----- 1630
Db 3144 E---QLSEKRVVAELK--SELAQTKLELETLKAQHKLKLEAFRLVYDKTDEVHLL 3198
QY 1631 -----SVREKPELTASTERVKNKMSNVVSGLPPEEFMLVYKARKHHT 1674
Db 3199 NDTLASEOKSRSELOWALEKEKAKLGRSEERDKEEL-----EDLKFSLESOKORNL 3249
QY 1675 TLTNLTETHTVYMKTDAEFVCERTL 1701
Db 3250 QL--NLLEBOOKOLNBSOOKIESORML 3275

```

```

RESULT 11
US-10-171-311-8
; Sequence 8: Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glaty, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoersch, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 3917
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-171-311-8

Query Match 3.5%; Score 333; DB 9; Length 3917;
Best Local Similarity 18.5%; Pred. No. 4.7e-09;
Matches 393; Conservative 337; Mismatches 775; Indels 622; Gaps 93;

QY 7 RVEEYQVNYNAMQKILCEPICLEIKPEVYSTRKDHIFKCFMGLNOKKSPSCPLCKN 66
Db 1339 QVOELESLSLQ-----QDLKTEQNTVEAFIHLQKRLQAVSESTVPSPSPVDSV 1389

```

```

QY 67 DITRSLQESTRETSOLVEELIKIICAFOLDTGLEVANSYNFAKKENN-----SPEHL 118
Db 1390 VITESOQRTMYGSCCKK-----NIDGTIEFSEFG-VKEETINIVKLEKQOEOL 1440
QY 119 KDEVS-1IOSM-----GYNRRA---KRLOSPEPNLSOETSLSVOLSNL 159
Db 1441 EEEVAKYIVMSJAFAQOETLSRISGGKENTASSKOHAAVCOOEOHAFNEMKLSODQIGF 1500
QY 160 GYVRT--LRTKORIOPOKTSVYTEL-----SDSEEDIVNKATYCSVDQDELO- 206
Db 1501 QTEETDVAKKEFFPLSK-----ELGEHKEILLNSDPPHIDPESKOCVULTISEEMFSKD 1556
QY 207 ---ITPOGTREDEI---SLDSAKRACEFSETD-----VTNTEHQPSNDLTTTEKRAA 254
Db 1557 KTFIVQSHIDEIVSVMASRQMLNEDQLEDMRQELVQOEHQOATELLNQAHHROM 1616
QY 255 ERPEPEYQSSVSNLHVEPCGTN---THASSLOHENSLLITDRANVEAFECNKSOP 311
Db 1617 ERQREDOE---OLQEEIKRLNROLAORSSIDEN--LVSEERVLLELE---ALKOL 1666
QY 312 GLARSOHNRWAGSKETCNDRTPTTEKKVDLMDPLCEERKNKOKLPCSENPREDIEDVP 371
Db 1667 SL-----AGREKLDCELRNNSSTOTQ---NGNE--NOGEVEBOQTFKEKELDRKPEDVP 1713
QY 372 WITLNS---SIQKVN-----EMFSRDELLGSDSDSHGESESNKV- 409
Db 1714 PELLSNRYALOKANNRLKILLEVYKTTAAVEETIGR--HVGILD-RSSKQSSASLI 1770
QY 410 ---ADVLDVINE-----VDREYSGS----- 425
Db 1771 WRSEAEASVSCVHEETRTVTDIESISYSGSDMPRNDINMMSKYTEGETELSORLYRSGF 1830
QY 426 ---SEKIDLLAS----- 437
Db 1831 AGTEIDPENNELMLNTSSRLQAAVEKLEAISETSOLEHAKVTOELMRSESRQOEAT 1890
QY 438 EALICKS---ERYHKS-----VESNIEDK-IFGKTYRKASLPLNLSHV 477
Db 1891 ESJLKQOEHLERLHEESRAREQALVELSKAEGVIDGVADEKTLFEPRIQOKTOI--IDRL 1948
QY 478 TENLIIGA---FTEPOIIOERPLTNKLKRRPTSG-----LHPEDFIKKADLAV 525
Db 1949 EOEELTASNRLOLEAEQOQIOEERELLSROKEMAAEAGPVEQOULOETEKLMKELEEV 2008
QY 526 OKTPREMIINGTONEQNGOVNITNSGHEKTKGDSIONKKNPNPIESLEKESAFPTKAE 585
Db 2009 OQCAEYVRDDLOK---OYKALEIDVEQVSRFTELEOEKNTELMDLRQONOLEKQLE 2063
QY 586 PI-----SSSISNMLELNT-----HNSKAPKK--NRLRKSS 616
Db 2064 KMKRFLDEQAIDREHENDVPOEIOKLQOLKAVPRPFRPISHQTEVEQLAHLKEKTD 2123
QY 617 TRHIALELVVSNLSPKNTLEIJDSCSSSEIKKKKYIOMVVRHSRNIQLMEGKE--P 674
Db 2124 ---KCESELLSKE-----OLORDIOERNEIEIKLEFR--VRELEQALVEDRKHFG 2169
QY 675 ATGAKKSNKNEQTSKRHSDTPEPCLKLTNAPG-----SFTK 711
Db 2170 AVAKAPRLSELVLOQAEORDAIDRKEKEITNLEQLEOFRELENKKEVVOQLHMOLEIOK 2229
QY 712 CSNTSELKEFVNPSPREKEEKLK--TVKVSNAEDPKDMLSGE--RVLOTERSESSS 768
Db 2230 KESTTRRLQEOENKLFKDKMEKLGIAIKESDAMSTQDOHVLGKFAQIIO--EKVEIIDO 2288
QY 769 ISLVPGIDYGTOSISILE---VSTLGAKKATPEPNKCVS-----QCAEFENPKGLIHGC 818
Db 2289 LN-----BOVYTKLOOOLKITTDNKNVIEEKNELRIDLETOIECLMSDOE-----C 2332
QY 819 SKDNRRND-----TEGFKYPLGHEVNHRSRETSIEMEESLEDAQYLONTEFKSKRSOFA 870
Db 2333 VKRRREIEIQLMEVIEKLOEQL---ANIGOKTSMANHSLSSEEDISIKHOLDVYVLEKLA 2389
QY 871 LFSNPGNAEECATFSAHSGSLKQSPKVTPECEO-----KEENOGKESNISIKPVOT 922

```

```

Db 2390 LEOOVETANE-WTF-----MKNVLKETNFKNMQOLQELFSLKRESEVVEKIQSPENS 2442
QY 923 VNTIAGFPVYGOKDP-----VDNAKCSIKGSRFLCSQFQKNETGLITPNKHGLON 976
Db 2443 VNVAIDLH---SKDKPELEVLVLTEDALKSLLENQTYF--KSFENGSGSIINETRETLQ- 2495
QY 977 PYRIPPLFIKSPFKTKCKKNLE-----ENFEHSMSPREKGNINIPSTVSTIS 1027
Db 2496 ---LESTVSAKDELEIIOCYQIKMDQOGEIEMLOKRIYVNOIKYEEKVAALYSOIQ 2552
QY 1028 RNNIRE-----NYFKEASSNI-----NEVSGSTNEVGSSINEIGSSD----- 1065
Db 2553 LEAVQEVAKFCQDMQOTISSEPERTINQNLQNRDELDGSDISALTLRITSELSQVEMHT 2612
QY 1066 ---ENIQALG--RNRCPKLNAMRLGVLOPEYVKQSLPSNCKNHPRIKQOIEVYQIV 1120
Db 2613 SLILEKQVEIAEKVNLEKKEKLLLEL-----QKLEENKREKREKRRSPQDVEVL 2664
QY 1121 NTFPSPLISDNL---QPMGSSHASQVSEPPDDLLDDGETIKEDTSPFANDIKESSAV 1176
Db 2665 KT--TTLEFHSNEESGFNLEBALRABSVATKA-----ELASTYKKEKLOEEL 2711
QY 1177 FSKSVQKGLSRSPSPFTHTHLAGYRGAKKLESSEENLSDEDELPFOHLLFGKVN 1236
Db 2712 LVKETNMTSLQKDLQ--VRDHLAFA--KEKLSLEKEDETEVOGSKKACMEPLPIKLSKS 2769
QY 1237 IPSQSTRHSTVATNECLSKNTENULSLKNS-----LNDGSONVYIAKASO----- 1281
Db 2770 IASO---TDGTLKLISSNQOTPOLIVKNAGIOINLOSECSSEVEIILISQIEKIEKMQ 2824
QY 1282 ---EHHLSE-----ETKCSASLFSQSCSELDLANTNTOPOFLIGSS--KQ 1323
Db 2825 ELHAELTDMESHRISETTEPLKREHVAVOLKEBCTGLAVALQCLRSK--GSSITPE 2880
QY 1324 MRQOSEQVGLSKDELAVSDERGTOLENNQEOGSDMSNLGEAAGCSEET---SVSE 1380
Db 2881 LAHSDAVQ---TREICSSDS--GSDMGOGIYLTSHQGFDIASEGGESEFSAJOSFPK 2933
QY 1381 DCSGLSSQSDILTTQORDTYQOHMLIKLOQMAELBAVLBOHGSOP--SNSIPSIISDSSAL 1439
Db 2934 KINGL-----LRAVHNEGMO--VLSTIESPYSDGEDHSIOQVSEPMLEEKRAVINIYISL 2986
QY 1440 EDLRNPEOSTSEKAVILTSOKSE-----YPISONPEGLSADKREVSDDSS 1484
Db 2987 KDLITKMQLOREAEVIDSSQSHESFSDWRGELLALQOVLLEBSVLLAFRIELTALGT 3046
QY 1485 T-----SKNKEPVERSSPSKCPSLDDRWYMHSCSGSLONRYPPOEELI----- 1529
Db 3047 TDAVGLLNCLEQRIQEOGVYEQAMECLOKADR-----RSLSEIQAHLAHQNM 3094
QY 1530 -KVAVDEQOLESGRPHDLTETSYLPRODLEGPIYLESIGLSLSDPESPESDRAPESA 1588
Db 3095 GRKITLKEDESEKPSOELLEYNIOQO---SOMLEMOVELSS-----MKDRATELO 3143
QY 1589 RVGNIPSPSALVPOLKVAESAOSPAAHITTPAGYNAMEE----- 1630
Db 3144 E-----OLSEKMYVAELK--SELAQIKLELTTIKAOHKHILKELEAFLEVKXKTDVEYHL 3198
QY 1631 ---SVSREKPELTASTERYVKNKRMVSVSGLTPEEFMLVYFARKHAI 1674
Db 3199 NDTLASQOKKSRLQONALEKAKLGRSEERDEEL-----BDLKSLSLSQOQRNL 3249
QY 1675 TLNLTLEETTHVYKTDAEFVCEPRL 1701
Db 3250 QI--NLLEEOQOLNLESQOKIESQRM 3275

```

RESULT 12  
 US-10-171-311-2  
 ; Sequence 2, Application US/10171311  
 ; Publication No. US20030087270A1  
 ; GENERAL INFORMATION:

```

: APPLICANT: Schlegel, Robert
: APPLICANT: Chen, Yan
: APPLICANT: Zhao, Xumei
: APPLICANT: Monahan, John
: APPLICANT: Kamatkar, Shubhangi
: APPLICANT: Galt, Karen
: APPLICANT: Ganavarapu, Manjula
: APPLICANT: Hoersch, Sebastian
: TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
: TITLE OF INVENTION: OF CERVICAL CANCER
: FILE REFERENCE: MRI-035
: CURRENT APPLICATION NUMBER: US/10/171,311
: CURRENT FILING DATE: 2002-06-12
: PRIOR APPLICATION NUMBER: US 60/298,159
: PRIOR FILING DATE: 2001-06-13
: PRIOR APPLICATION NUMBER: US 60/298,155
: PRIOR FILING DATE: 2001-06-13
: PRIOR APPLICATION NUMBER: US 60/335,936
: NUMBER OF SEQ ID NOS: 238
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO: 2
: LENGTH: 3907
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-10-171-311-2

```

Query Match 3.4% Score 329; DB 9; Length 3907;

Best Local Similarity 18.5%; Pred. No. 7.6e-09;

Matches 367; Conservative 313; Mismatches 726; Indels 576; Gaps 89;

```

QY 7 RVEYQVNVANQKILCEPICLELKEPVSTKCDHIFCKFCMLKLNQKGPSCPLCKN 66
DB 1339 QVQESLSSIQ-----QOLKETQNEAEIHCLOKRLQAVSESTVPSSLVDVY 1389
QY 67 DITKSLSQESTRFSQVLEELKILICAFQDITGLEVANSINFAKKN-----SPEHL 118
DB 1390 VTTEDAOQNTWPGSCVKK-----NIDGTIEFGEGF-VKEETNIYKLEKQYOEOI 1440
QY 119 KDEVS-IIOSM-----GYKNRA-----KRLQSPENPSLOETSLSYQSLN 159
DB 1441 EEEVAKVIVSMIAFAQOELSRISGKENTASSKQAAVCOEOHYFNMKILSQOIGF 1500
QY 160 GTVRT--LRTQRIOPKTSYVIELG-----SDSEEDTVNKATYCSVQDOELLO 206
DB 1501 QFFETVDVYFKFEPFLSK-----ELGEHKEILLNSDPHDIPESDCVLTISEEMFSD 1556
QY 207 ---TTPQGRDEI---SLDSAKKACEFSETD-----VTNTEHHOPSNNDLTTEKRA 254
DB 1557 KTFIVRQSIHDEIVSSMDASQMLNEBQLEDMROELVROYOEHQOATELLRQAHMRQ 1616
QY 255 EHPPEKQGVSSVNLHVEPCGTN---THASSLOHENSLLTPKDRNNVKAFCNCKSKOP 311
DB 1617 ERQREDOE-----QLOEITKRLNROLAQSSTIDNEN--LYSERERVLLELE---ALKOL 1666
QY 312 GLARSQHNWAGSKETCNDRTPTSTEKYVDLADPLCERKENKOKLPCSENPRTDPEY 371
DB 1667 SL-----AGREKLCELRNNSQTQ---NGNE---NOGEVEQYFFKEELDRKPEDYV 1713
QY 372 WITLNS---STQKVN-----EMFSRDELLGSDSDHDESESNAKV- 409
DB 1714 PEILSNERVALKNNRLKILLEVYKTTAAVEETIGR--HYLGILD-RSSKSSQSSASLI 1770
QY 410 -----ADVLVDLNE-----VDEYSGSS---EKIDLLASDPHEALICKSERVHS-- 449
DB 1771 WRSEADASYKSCVHEHRTVDESTIPSYSGDMPRDINMWSKYTEEGELSORLYRSGF 1830
QY 450 --KSVESNIEDKIFGKYRKAKSLPNLSHV---TENLIGAEVTEPOII-----QERPLT 499
DB 1831 AGTEIDPENEELMLNISSRLQAAYEKLLEAISSETSSQLEHAKVYQTELMRESFRQOENT 1890
QY 500 NKLKRRKRTSGLHPED-----FIKAD-----LAVOKT--PEMINOGT---NOTEO 541

```

```

DB 1891 ESLKQCEELRERLHESRARQELAVELSKAEVIDGYADEKTLFEROIQOENTDIIIDRLQ 1950
QY 542 NGQVNNITNSGHNKTKGSIONEKNPNPIELESKSAFKTAEP-----SSISNM 594
DB 1951 --ELCASNRLQELAEQOIOEBERE---LISROKE-AMKAGVYEQOOLQOETEKLMKE 2004
QY 595 EELNINHSKAPKNNLRKRSSTRHIALELVVSRNLSPPNCTELOIDSCSSSEIKKK 654
DB 2005 KLEVOQAKV-----RDDIQOYKALEIDVEQVS----- 2035
QY 655 YNOMPVRHSRNQLMBGKEPATGAKSNK-----NEOT-SKRHSDTFP-----ELKL 702
DB 2036 -REIELEQKNTLMDLRQONALEQOLEKMKRFLDEQALIREHREDVFOEOIKLEOOL 2094
QY 703 TNAPGSEFKCS--NINSELKEFVNPSLPREKKEKLETKVSVNNADPDLMLSGRYLOT 760
DB 2095 KVPV-RFQPISEHOREVQOLAN---HLKEKTDKSEILLKSE-OLORDIOERNEDEIKL 2149
QY 761 ERSVESSISLVPGDYGTOESISLLEVSTLGAKTEPNKCVSQCACAFENPKGLHGC 820
DB 2150 EFRVRELEQALLVSAD--TFQKVE--DRKHFGAVAKPE--LSLEVQOARDAIDRKEX 2203
QY 821 DNRNDTEGFKYPLGHEVNHRSRETSTIEMBESELDQOYLONTERKVSROSFALPSN 880
DB 2204 EITN-----LEQLEQFRE---ELEKNKEEVOQLHMOLEIOKKESTTRLO--ELEQ 2249
QY 881 ECATES-----AHSGSLKOSPKVTFE-----QEOKE-ENOGNESINIPVQTVN 924
DB 2250 ENKLFQDDEKLGALAKESDASTODQVHLFKFKQIITOEKREVELDQI-NEOVTYKLOOOL 2309
QY 925 ITAGFPVQKDKPVNNAKCSIKGSRFCLSSQFNGNETGLTPPKHGLONPYRIPLF 984
DB 2310 ITDNKVIEKKNELIIDLETOIE---CLMSD----- 2337
QY 985 PIKSPFKTKCKNLLENFEE-----HSMSPER----- 2402
DB 2338 -----QECVARNREELTEQLENEVIEKLOELANIGQTSNMASHISEEDSLKHOLDV 2390
QY 1013 -----EMGENIPSTVSTISRNINENFK-----EASSNINEGS 2449
DB 2391 VIAEKALQOYETANEEM-----TFKNVILKETFMKNOLTOELFSLKRESVEKIQS 2445
QY 1050 -STNEYGSSINIGSSDENIQALGRN-----RGFKLAMRLVGLQ 1090
DB 2446 IPENSVNVAIDHLSKDKPELEVLTEDALKSLQENQTFKSPBENGKGSITINLFTRLLO 2505
QY 1091 PEYTKOSLPGSNCKHPEIKQYEEVQVNTDFSPYLSIDNLEQPMGSSHASQVCESTP 1150
DB 2506 STVSAKDELDTQC-YKQIKRMOEQOQFETEMLOKKIVMLQKIVEKVAALVSOIOLEA- 2563
QY 1151 DDLIDGFEIKEDTSPAENDIKESSAVFSKSVQ-----KGLERSPSPTF----- 1194
DB 2564 -----VOETAKCQDNQNTISSPERTNIONLNQLEDELGDSIALTLRISELSQV 2615
QY 1195 ---HTHLAAGYRRGAKLLESSEENLSSEDELPQFQHLIFGKVVNNIPDSQSTRHSTVATEC 1251
DB 2616 VEMHTSLI---LEKEQVEIAEKVLEKEKKLLELOKLEBNEKKORKEKRRPOVEY 2671
QY 1252 LSKNTEENILSKNSLNCDSNOVILAKASQEHILHSEFKKCSALFSSQCSLEBLUTANTN 1311
DB 2672 LKTTTE---LFHNSNEESGFNLEALRAE---SVATKABLASYKKAEXLEQLELLVKE 2723
QY 1312 TQDPELIGSKQMR-HQSES--QGVALSKELEVSDEER----- 1347
DB 2724 TMTSLQKDLQSVRHLHALEAKKSLILEKEDTEVQESKAKCMPEPLPIKLSKIASOTD 2783
QY 1348 GT-GLEENNQOEQSDSNLG--EASQCESETSVEDCSGLSSQ-----SDILTT 1394
DB 2784 GTLKISSNQTPQILVKNAGIQINQSCSESE-EVTEIISQFTKIEKMQELHAAETLDM 2842
QY 1395 QORDTMQHNLIK-----LOQEMALBLAVLEQHGQSPNSPISII-SUSSALEDLRN 1444

```

Db 2843 ESRHISTETFLKRBHYAVOLLEECGTAKAVIQCRLSKRGSSIPETLAHSDAYQTRICIS 2902  
QY 1445 PEOSTSEKAVLTQSSESEYIPSONPEGLADKEFEVSADSTSKNKEPVERSSPKCPSL 1504  
Db 2903 SDGSGMDGOGIYLTHSGFPIASEGRG---ESESANDSPKKIK---GLIRA----- 2949  
QY 1505 DDERWMSHSGSGLONRNRYPSQOEELIKVVDYEQOLESFGPHDLJETSYPRLQDEGTPYL 1564  
Db 2950 ---YHN-----EGMOVLSTLESPPSGEDHSIQOVS-----EPWL 2981  
QY 1565 ESGISLFSDDPESDPSDRAPESARVGNIPSSSALKVPOL---KYAESASQSP----- 1614  
Db 2982 EE-----RKAYINTLSTLKDILTKMQLQREAEVYDSSQSHESFSDW 3022  
QY 1615 ---AAHPTDTAG---YNAEESYRREKPELTASTERY 1646  
Db 3023 RGEILLALQOVFLERESVLLAARTELTLALGTIDANGLLNCLQRLOEGOVEYQOAMECL 3082  
QY 1647 NK 1648  
Db 3083 OK 3084

RESULT 13  
US-10-171-311-6  
Sequence 6, Application US/10171311  
Publication No. US20030087270A1  
GENERAL INFORMATION:  
APPLICANT: Schlegel, Robert  
APPLICANT: Chen, Yan  
APPLICANT: Zhao, Xumei  
APPLICANT: Monahan, John  
APPLICANT: Kamatkar, Shubhangi  
APPLICANT: Ghatt, Karen  
APPLICANT: Gannavarepu, Manjula  
APPLICANT: Hoeish, Sebastian  
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY  
TITLE OF INVENTION: OF CERVICAL CANCER  
FILE REFERENCE: MRI-035  
CURRENT APPLICATION NUMBER: US/10/171, 311  
CURRENT FILING DATE: 2002-06-12  
PRIOR APPLICATION NUMBER: US 60/298,159  
PRIOR FILING DATE: 2001-06-13  
PRIOR APPLICATION NUMBER: US 60/298,155  
PRIOR FILING DATE: 2001-06-13  
PRIOR APPLICATION NUMBER: US 60/335,936  
PRIOR FILING DATE: 2001-11-14  
NUMBER OF SEQ ID NOS: 238  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 6  
LENGTH: 3925  
TYPE: PRN  
ORGANISM: Homo sapiens  
US-10-171-311-6

Query Match 3.4%; Score 329; DB 9; Length 3925;  
Best Local Similarity 18.5%; Pred. No. 7.7e-09;  
Matches 367; Conservative 313; Mismatches 726; Indels 576; Gaps 89;

QY 7 RBEVONVINAMOKILECPICLLELIPVSTKCDHIFCKEMLKLNOKKGPSCPLCKN 66  
Db 1339 OVQELSLSSLD-----QOLKETQONVFAELHCLQKRLQAVSESVPLSLVDV 1389  
QY 67 DIRKSLQESTRSQVLBELKLTICAFOLDTGLGTANSTYFAKKNN-----SPEHL 118  
Db 1390 VITEESAQRTMPYGCVK-----NTDGTIEFGSEFG-VKEETNIVKLEKYOYEOEL 1440  
QY 119 KDEVS-IIGSM-----GYRRA---KRLQSEPNLSLOETLSVQLSNL 159  
Db 1441 EEEVAAVYMSIAFAQOETELSRISGKEMTASSKQAHAVQOQOYHFNMKLSQOIGF 1500  
QY 160 GTVRT--LRTKRIOPQKTSVYIELG-----SDSESDTVNKAITYCSVDQOELLQ- 206

Db 1501 QTEETVDVKKFEKEFKPLSK-----ELGHEGKEILLNSDPPHIDIESKCVLTISEMFSKD 1556  
QY 207 ---ITPGTDEI---SLDSAKKACFESETD-----YNTTEHOPSNNDLNTTERAA 254  
Db 1557 KTFIVROSHIDELISVSSMDASROMLNEQLEDMROELVAYOYQHOQATELROAHNRQ 1616  
QY 255 ERHPEKYQGSVSNLHVEPCGTN---THASSLOHENSLLTFRDMRVAEFCNSKQSP 311  
Db 1617 ERQREDE-----QLOEIKRLNRQLAQRSSIDEN---LVSEBRVLELELE---ALKOL 1666  
QY 312 GLARSOHNRWAGSEKTCNDRPTSTEXKVDLNDPLCEKEMWKOKLPCSEMPREDVDP 371  
Db 1667 SL-----AGREKLCCELRNLSSTQO---NGNE---NOGEVEQTEKEKELDRKPDVP 1713  
QY 372 WITLNS---SIOKVN-----EMFSRDELLGSDSHGSESNAKY- 409  
Db 1714 PETLSNERVALQKANNRLKILLEVYKTTAABETIGR---HVIGILD-RSSKQSSASLI 1770  
QY 410 ---ADVLDVINE-----VDEYSGSS---EKIDLIASDPHALICKSRVHS-- 449  
Db 1771 WRSEAEASVKSCHYEHTRTVDESIPSYSGSDMPRNDINMWSKYTEGETLSQRLVRSGF 1830  
QY 450 ---KSVESNIEDKJFGKTYRRKASLPNLSHV---TENLIGAFYTEPOI---QERPLT 499  
Db 1831 AGTEIDPENELMLNLTSSRLQAAVEKLLAISETSQLEHAQVYQTELMRESFROQEAR 1890  
QY 500 NKLKRRRPTSGIHPED-----FIKAD-----LAVKT---PEINNGT---NOTEQ 541  
Db 1891 ESUKCOEELRERHEESRAREQALAVELSKAEGYIDGVADEKTLFEROIQERTIIDLEQ 1950  
QY 542 NGOVNNTNGHINKTKGDSIONEKNPNPIESLEKESAFYTKAEPi-----SSSISM 594  
Db 1951 ---ELUCASNRLQLELAQOQIOERE---LLSQKE-AMARAGPYQOQLQOETEKLMKE 2004  
QY 595 ELELNTHNSKAPKKNRLRRKSTRHIALELVVSRLNPPNCTELOIDSCSSSEETKRRK 654  
Db 2005 KLEVOQCAEKV-----RDLQKOVKALIDEVQVS----- 2035  
QY 655 YNMPVHSHNLDLMEKEGATGAKSKNR-----NQGT-SKRNDSTFP-----ELKL 702  
Db 2036 -RTELEQERNTLMDLRQONQALEKQLEKRRFLQOALDRHERVFOETLOKLEQOL 2094  
QY 703 TNAPGSFTKCS---NTSELKEFVNPSPREKEEKELETVKVSNADEPKDMLSEVLOT 760  
Db 2095 KVPV-RFQPISEHOTREVEQLAN---HLKERTKCSLLSKF-QLOORDQERNEETKL 2149  
QY 761 ERSVSSSISLVPGTDTYQOESISLLEVSTLGRAKTEPNKCVSOCAFENPKGLIHGCSK 820  
Db 2150 EFRVRELEQALLVASD--TFQVKE--DRKHFGAVEAKPE--LSLEVOQLQERBALDREK 2203  
QY 821 DNRNDTEGEKPYPLGHEVYNSRETSIEMESELDAQVLYONTFKYSKROSFALFSPGNAEE 880  
Db 2204 EITN-----LEEOLEQFRE---ELEKNNEEVOQLHMOLEOKKSTTRLO---ELEQ 2249  
QY 881 ECATFS-----AHSGLKQSPKPYFE-----CEQKE-ENGKNESNIKPYQVTN 924  
Db 2250 ENKLFKDDMKDLALAKESDAMSTODOHVLFGFAQIQOKEVEIINDLQVYTKLOOQOLK 2309  
QY 925 ITAGFPVYQKDPVDNAKCSIKGSRFCULSOFNGNETGLIPNKHGLQNPRIPLP 984  
Db 2310 ITDNNKVIIEKKNLIDLEQOIE-----CLMSD----- 2337  
QY 985 PIKSPYTKCKKMLLENPE-----HSMSPR----- 1012  
Db 2338 ---QECVKNRREEEIQLNEVIEKLOQELANIGOKTSMNAHSLSEADSLKHQOLDV 2390  
QY 1013 ---MGNENIPSTVSTISRNIRENVK-----EASSSINVEGS 1049  
Db 2391 VIAEKALAEQVTTANDEM-----TMKNVYLKETNFKMNOJLOELSLKNERBSVERIQS 2445  
QY 1050 -STNEVGSSINTEIGSSDENIOAELGRN-----RQPKINAMLRGLVIO 1090

```

Db 2446 IPENSVAIDHLSKDKPELVLTEDALKSLENOTYFKSFEENKGSIIINLETRLOLE 2505
QY 1091 PEVYQSLPGSNCKHEIKKOEVEVVOYVNDPFSYLLSDNLQOMGSSHASOVSETP 1150
Db 2506 STVSAKMDLELTQC-KYQIKDQBOGOFETEMLOKKVYNLOKTYEEVVAALVSQIOLEA- 2563
QY 1151 DDLDDGEIKEDTSPFAENDIKESSAVFSKVQ-----KGELSRSPPT----- 1194
Db 2564 -----VQEVAKPCQDNQOTISSEPERINIQNLNLDREDELQSDISALTRISELESQV 2615
QY 1195 ---HTHLAGYTRGAKKLESSEBNLSDEBELPCFOHLLFGKYNNIPSOSTRSTVATEC 1251
Db 2616 VEMHTSLI-----LEKEQVEIAEKVLEKELKLELOKLEGNKKKOREKERRSPDVEV 2671
QY 1252 LSKNTEENLSTLKNLSNDCSNQVITLAKASQEHHLSPETCSALFSQSCSELDLANTN 1311
Db 2672 LKTTTE---LPHSNESGFFNEELARAE-----SVATRAELASVKEKAEKLOEELLVKE 2723
QY 1312 TQDPELLIGSSKOMR-HOSES-QGVGLSDKELVSDDEER----- 1347
Db 2724 TMTVSLQKDLQVROHLAEAKELSLKEKDETEVQESKKACMFEPPLIKLSKSIASQTD 2783
QY 1348 GT-GLEENNOEQSMDSNIG---EASGCESETSVESDCSGSSQ-----SDILTT 1394
Db 2784 GTLKISSNQOTPOILYKNAGIOINLOSECSSE-EVTEIISOFTKEIKEMOELHAAEILDM 2842
QY 1395 QORDMOMHLIK-----LOQEMAELEAVLEQHSQSPNSPYSLI--SDSSALDLRN 1444
Db 2843 ESRHSETETLKRREHVAVOLKRECGITAKAVIQCRSKSEGSIPELANSDAQOTRELC 2902
QY 1445 PEQSTSEKAVLTSOKSSEVPISQNPGLSADKEFVSADSTSKNKEPVERSSPKCPSL 1504
Db 2903 SDGSDMGQGYLTHSGQFDIASSEGRG---ESESATVDSFPKKIK--GLLRA----- 2949
QY 1505 DDMWMSGSSGLQNRNYSQOELIKVYVEEQOLEESGPHDLTETSYLPRLQLETPYL 1564
Db 2950 -----VHN-----EGMOVLSTLSPSYSDGHDHSIQVYS-----EPWL 2981
QY 1565 ESGISLFDSPDSEDSEDRAPESARGNIPSSTSALKVPOL-----KVAESAQSP----- 1614
Db 2982 EE-----RKAVINTISSLKDLITTKMOLOREALVYSSOSHEFSQDM 3022
QY 1615 -----AAHTTDTAG-YNAMEESVREKPELTASTERY 1646
Db 3023 RGEILLALQVYLEERSVLLAARTELTATGTTDAVGLNLCLEQRIQOEGVEYOAMECL 3082
QY 1647 NK 1648
Db 3083 QK 3084

RESULT 14
US-09-815-242-5635
; Sequence 5635, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23

```

```

; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5635
; LENGTH: 2368
; TYPE: PRF
; ORGANISM: Staphylococcus aureus
US-09-815-242-5635

Query Match      3.4%; Score 328.5; DB 10; Length 2368;
Best Local Similarity 18.2%; Pred. No. 4e-09;
Matches 349; Conservative 306; Mismatches 720; Indels 547; Gaps 78;

QY 93 FOLDGLEYANSYNFAKKENNSPEHLKDEVSIIQSGYRNRA-----KRLQSE 141
Db 390 FTINT--EIGNNGNFQ--ASLKADQFKYEVTLPQGVTVYVNSLTTFPNGNEDSTVLKNM 445
QY 142 PEPNLSQETSLSVOLSNLCTVTLTKRIQIQPKT---SVYIELGSDSEDTVN---KAT 195
Db 446 TVNYDQANAKVFTSGQVTTARGTHTKEVLEFPDKSLKLSYKVVANANIDTPKNIDFNEELT 505
QY 196 YCSVGDOELQITPOGTREIISLSAK--KAACEFSETDVNTNHHOPSNNDLNTTEKRAA 254
Db 506 YRRASD-----IVINNAQEVITLADPESVAENMKKDALQOQVNSQVNSHYTTAS--IA 558
QY 255 ERHPEYQGSVSN--LHVEPC-----GTYTHASSLOHENSILLTKRRNNEKA 302
Db 559 EYKRLKQADNITLNEBANDHVEFANRASAQALDGLVTKLQALIDMQAALAEIDAKAOKEV 618
QY 303 BECNKSKQGLARSOHNRWAGSKETCNDRPSTSEKVKVDLANPCEKKE-----MWNQK 357
Db 619 TAAQGSKKV-----TQDEVALVATKINNDKNNAAIAELNKQTTAAGYTTKDKNGIAYLDQV 674
QY 358 LPCSENPRTEDVPWITLNSIIOKVNEWFSRSDLLGSDSHGSESNKAVADVLDVYL 417
Db 675 ITPTVAPQAKQDI-----IQAVT---TRQOI-----KKSNAILOQEKDVAN 713
QY 418 EYDEYSGSSEKIDILASDPHEALICKSERVHSKSVESNIEDKIGKTYRKASAPNLSHV 477
Db 714 D-----KIGKIEYKAIKIDDAITNAQVEALKTAKINDIQATPATYAKAAALEEEDV 767
QY 478 TENLI-----IGAFVTEPOITIOERPLNKLKRRRPTSGLHPEDFIKKADLAVQKTPKIN 533
Db 768 VOQOIDQAPLNDPTTNEEVAEAIERIKNAK-----VSGVK-----AIEATTTAOD 812
QY 534 OCTNQTEQNGVAMNTTNSCHENKTKGDSITONEKNPNPIESLEKSAFKAEPKAPISSSISN 593
Db 813 LERVKNEELSKIENTDS---TQTKMDAYNEVK-----QADARTKQATVATYSNMTNE 861
QY 594 MELELINHSKAPKKNRLKRSSTRHITALELVSR-----NLS 632
Db 862 EVAEADAVEAOKOG-----LHDIOVYVSKQEVADTKSKVLDKINAIOQOAKV 911
QY 633 PNCNT-----ELQIOSCSSEEEKKKKYQOMVYRHR---NLQMEGEPATGA 678
Db 912 PAADTEVENAYVTRKQEOIENSNASTTEE-KQATVTELTQKQOEARNTIDAANTSDVYTA 970
QY 679 KKS-----NKPEQTSKRHSDTPPELKLITNAPGSFTKCSNTSELKEFVNPSLPREK-- 731
Db 971 KONGIAALNQOAAATTK--SDAKAEIA-----QASBRKTAIEMNDSTTEEDQAA 1020
QY 732 EKKLETIVKYSNNAEDPKDMLSGERYLOTEBSVSSSISLVPDGYGTQSSISLFE---- 787
Db 1021 KDKVDAQAVVTANAD--IDNAAAMTQVDNAKTTNEATITAIATPDPANVAPTAQAIAIDKVOA 1078

```





```

QY 358 LPCSENPRTDEVPWITLNSIOAKVNEWFSRDELLGSDSDHGESESNAAVADVLN 417
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 675 ITPYKPOAKODI-----IQAVT-----TRKOI-----KKSASIODEADVAN 713
QY 418 EVDYSGSEKIDLLASPHEALICKSERVHKSVSUESNEDIKFGTYRKKSLEPNLSHV 477
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 714 D-----KIGKIEFAIKDIDAATTNAQVEAIKTKAINDINOTAPATTKAALIEFDEY 767
QY 478 TENI-----IGAFVTEPOLIOERPLTNLKKRRRPTSGHPEDFIKKALADQKPEMIN 533
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 768 VOAOIDAPLPDPTNEEVAEIERINAAK-----VSGVK-----ALEATTAOD 812
QY 534 QGTNOEONGOVMTNITNSGHEKTKGDSIONEKNPNPIESLEKESAFKTAPESSISN 593
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 813 LERKNEISIKIENTIDS---TOTKMDAYNEVK-----QAAARKQONTVSATNE 861
QY 594 MELEINTNSKAPKKNRLRKSSSTRHIALLEYSR-----NLS 632
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 862 EVAADAAVEAAOKG-----LHDIOYVKSKEOVADTKSKVLDKINAIOQAQVK 911
QY 633 PPNCT-----ELQIDSCSSSEIKKKKYOMPVHRSR-----NLQMGKEPATGA 678
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 912 PAADTEVNAYNTRKQEIQNSNASTIEE-KQAAITELDTKKOEARTNLDAANTSDVYTA 970
QY 679 KKS-----NKPNQTSKRHSDTFPELKLITNAPGSETKCSNTSELKEFPNPLPREK-- 731
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 971 KDNGIAAINOYOAATTKK--SDAKAIEA-----OKASERKTAIEAMNDSTEBQQA 1020
QY 732 EEKLETYVSNNAEPKDLMSGERVLOTESVSESSISLVGTQVGOESISLE---- 787
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1021 KDKDQAVVYTNAD--IDNAAANTOVDNAKTTNEATIAITPDANVKPTAKAIDKQYA 1078
QY 788 -----VSTLGKAKTEPNKCVSOCAFEENPKGLIHGCSKDRNDETEGKPYLGHEVNSRPT 843
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1079 QETAIIDANGATTEKAAKQOVTEK-----TTADTAID-----GAHTAEVYA 1123
QY 844 SIEMESELDAQYLONTEKVSROSFALFSNPG-----MAEECAPTESAHGSLK 893
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1124 AKNEIAIKIEAIQPATTTKDKNAKQAIATKANERKTAIAQOTIDITAEIAMAANVDNAV 1183
QY 894 KQSKRYTECEQKEENOCK--NESNIKPY-QTY-----NITAGFVVGOKDKP 938
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1184 QANNNIEAANSONDVQAKTTGEASIDQVTPVNKKATAVTDAKNNITA-----ATDNG 1238
QY 939 VDNKK-----CSIKGSR-----FCLSOFRGNETGLITPNKHGL----- 973
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1239 VDTAKDAGKNSIQSTQPAFAKSNKANDVDQAVTQONQALDNTTGATTEKNAADVLK 1298
QY 974 -----LONPYRIPLLEPIK-----SPVYTKCKKNLLEENFEHSM--- 1008
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1299 AKERAYODILNAQTTNDVTOIKDOAVADVOGITADTTIKDVAKDELATKAREOKALIAQT 1358
QY 1009 -----SPERE-----MGNEPIPTVSTISRNNI RENVFKEASSNINEVGSSTN- 1052
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1359 ADATTEKEQANOQVDAELTQONQNIENAOSTIDVYTAQDNALQA-----IDPIQASTDV 1413
QY 1053 -----EVGSSINEIGSSDENIQAELGRNRP-----KLNAMLRG-----V 1088
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1414 KTNRAELLETMOUKITELILNNETTNEKGNIDIGVRAAYEEGLNINIAATTTGDVYTA 1473
QY 1089 LOPEYVK-QSLPGSNCKHPEIK-----KOEYEEVOQVNTDPSF 1126
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1474 KDTAVQVQOQLHANPVKKPAGKTAALDQAADKKTQIEQTPNASQOEINDAKOEVDTELNO 1533
QY 1127 YLIDNLEQPMGSSIASOVCEPDDLLDGE-----IKEDTSFAENDIKESSAVFSKY 1181
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1534 --AKTNIDQ-----SSTDEYVDNAVKEGAKINAVKTFSEYKRDALAKIEAAYNAKV 1583
QY 1182 OKGELSRSPDFTHTHLAOGYRGAKKLESSEPL-----SSEDEELPCFOHL----- 1229
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1584 TEADNSNAA--TSSEIAEKOKLAEKOTADQONQATSKDDIEVOIHNDIDNINDYTI 1640
QY 1230 -----LFG-----KVNNIPSOSTR--HSTVATECLSKNTEENLILKNSIND-- 1269

```

```

Db 1641 PTGKRESATTDLYAYADQKKNNISADTNATODEKQOAIKQYDQONQVTALESINNGVNDG 1700
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1270 -----CSNOYILAKASEHHLSEETCSAALFSSQSELEDL 1306
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1701 VDDALTOGKAAIDAIOVDATVKPKANQVYIDAKA-----BETK-----ESIDQSDQL 1746
QY 1307 TANTNTODPFLIGSSKOMHOSQGVLSPEKEVSDDEE-RGNGLE-----ENNQOE 1358
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1747 TABEKTALAMT---KQITDOAKO---GITDATTTAEVEKKAQGLEAFNDIOIDSTEQ 1800
QY 1359 QSMDSNLGFAASCESETSVSEDCSGLSSQSDILTTQORDTMQHL-----IKLO 1408
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1801 KATIE-----ELETAIDQIEAGVNVNAD-ATTEKEKAFTALIEDILSKATEDISDQ 1849
QY 1409 QEMAELEAV-----LEOHQSOPNSYPSIISDSSALEDRNEQOSTSEKAVLTSQSSSEY 1464
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1850 TTNAEIATVKNLSALBOLAQRINP---VVKNALEAIREVYNNQIEIKNADADASKE 1905
QY 1465 ISONPEGLSADKFEVSADSTSKNE-----PVERSSPSKCPSLDDRMYMHCSCG 1515
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1906 IARTDLGRYDFRFAKLD-KTQTNTEVAEQLQVNTTIPALEIYVPONDPRAND-----TNSG 1959
QY 1516 S-----LONRNPSOEELIKVVDY----- 1534
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1960 SONMDATANSANATAPENTGPNVTESTDNANADTSSTTNONDAATGETTATSANSSA 2019
QY 1535 -----EEQLEESGPHDLIETSYLPRODLEGPYLESGISLFSDDPESDPEGDAPSAR 1589
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2020 TODDANDKPOANNSSADTSTNSPTMDNDVTSKPEVESTNNCTTDPKATEADNATPAESAT 2079
QY 1590 VGNIPSSSALKVPOLKVAESAOSPAHAHTTDTAGYNAMESV-----SREKPELTAS 1642
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2080 NNNSTTTATNENAP-----TESTATAPPTASTGAESSADSKONASVNDSKQAAEVNNS 2132
QY 1643 TE 1644
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2133 AE 2134

```

Search completed: June 27, 2003, 10:27:35  
Job time : 68 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 27, 2003, 10:13:36 ; Search time 62 seconds  
(Without alignments)  
2888.684 Million cell updates/sec

Title: US-09-734-672-4  
Perfect score: 9649  
Sequence: 1 MDLSALRVEEVQNVINAMQK.....LYOCQELDTYLIPOIPHSY 1863

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_73:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9642	99.9	1863	1 A58881	breast/ovarian can
2	4882	50.6	1812	2 T49350	breast/ovarian can
3	373	3.9	2447	2 T16870	hypothetical prote
4	365.5	3.8	3329	2 T42205	breast cancer susc
5	361.5	3.7	3343	2 T42207	breast cancer susc
6	358.5	3.7	5327	2 T13564	microtubule-associ
7	357.5	3.7	3329	2 T30904	breast cancer tumo
8	355	3.7	2954	2 T14156	kinesin-related pr
9	342	3.5	2271	2 F90073	hypothetical prote
10	341.5	3.5	1837	2 T41023	probable nuclear p
11	340.5	3.5	1837	2 JC5837	364K Golgi complex
12	340	3.5	2829	2 A42771	reticulocyte-bindi
13	339.5	3.5	3488	2 T34418	hypothetical prote
14	339	3.5	1957	2 T38077	hypothetical colle
15	334.5	3.5	3328	2 T30835	breast cancer tumo
16	333	3.5	6713	2 B89921	hypothetical prote
17	332	3.4	3225	2 T52300	giantin - human
18	330.5	3.4	3147	2 T18674	hypothetical prote
19	328.5	3.4	5170	2 T15348	hypothetical prote
20	327	3.4	1875	2 S38173	myosin-like protei
21	327	3.4	3418	1 G02354	giantin - human
22	323	3.3	3259	1 A56539	breast cancer tumo
23	322	3.3	1658	2 S55101	hypothetical prote
24	321.5	3.3	4377	2 A55575	ankyrin 3, long sp
25	320	3.3	1738	2 T14867	interaplin - slime
26	319	3.3	65	2 G02989	BRCA1 - Thesus mac
27	319	3.3	1410	1 A57013	early endosome ant
28	318.5	3.3	1526	2 A45605	mature-parasite-in
29	317.5	3.3	3507	2 T34513	hypothetical prote

30	317	3.3	2346	2 T13829	Tpr homolog - fru1
31	316.5	3.3	2484	2 T26216	hypothetical prote
32	316.5	3.3	2607	2 T26215	hypothetical prote
33	316	3.3	1727	2 T50073	myosin-like coiled
34	315	3.3	2481	2 D90011	FMb protein (limp
35	314	3.3	1871	2 D96796	probable heat shoc
36	313.5	3.2	1679	2 S48385	hypothetical prote
37	312.5	3.2	1790	2 S67593	transport protein
38	312	3.2	2938	2 T30249	cell proliferation
39	312	3.2	3924	2 S37431	ankyrin 2, neurona
40	311.5	3.2	2748	2 S57976	nuclear migration
41	310.5	3.2	2845	2 T49505	adenomatous polypos
42	309.5	3.2	2253	2 T30336	nuclear/mitotic ap
43	304.5	3.2	1744	2 JH0720	tanabin - African
44	301	3.1	1269	2 F84730	probable myosin he
45	301	3.1	2464	1 QMSP1	microtubule-associ

## ALIGNMENTS

RESULT 1  
A58881  
breast/ovarian cancer susceptibility protein BRCA1 - human  
C:Species: Homo sapiens (man)  
C:Date: 16-Oct-1998 #sequence-revision 16-Oct-1998 #text-change 18-Jun-1999  
C:Accession: A58881; A54652; I58130  
R:SKOLNICK, M.H.  
submitted to GenBank, September 1994  
A:Description: Human breast and ovarian cancer susceptibility (BRCA1) mRNA, complete  
A:Reference number: A58881  
A:Accession: A58881  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-1863 <SNO>  
A:Cross-references: GB:U14680; NID:G555931; PIDN:AAA73985.1; PID:G555932  
R:MIKI, Y.; SWENSEN, J.; SHATTUCK-BIDENS, D.; PUTREAL, P.A.; HARSHMAN, K.; TAVLIGIAN, Y.E.; HATFIELD, T.; PHELPS, R.; HAUGEN-STRANO, A.; KATCHER, H.; YAKUMO, K.; GHOLAMI, OW, P.K.; MORRIS, F.H.; HELYERLING, L.; MORRISON, P.; ROSTECK, P.; LAI, M.; BARRETT, J Science 266, 66-71, 1994  
A:Authors: Lewis, C.; Neuhausen, S.; Cannon-Albright, L.; Goldgar, D.; Wiseman, R.; K A:Title: A strong candidate for the breast and ovarian cancer susceptibility gene BRC A:Reference number: A54652; MUID:95025896; PMID:7545954  
A:Accession: A54652  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-444, 'D', 446-487, 'S', 489-593, 595-600, 'W', 601-1452, 'VLQ', 1456-1471, 'X', 14 A:Cross-references: GB:U14680  
R:HOSKING, L.; TROWSDALE, J.; NICOLAI, H.; SOLOMON, E.; FOULKES, W.; STAMP, G.; SIGNE Nature Genet. 9, 343-344, 1995  
A:Title: A somatic BRCA1 mutation in an ovarian tumour.  
A:Reference number: I58130; MUID:95315980; PMID:7795636  
A:Accession: I58130  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1812-1813, 'PGQRTMASMOIGRCVRHLM' <HOS>  
A:Cross-references: GB:S78558; NID:G1000568; PIDN:AAAB34725.1; PID:G1000566  
A>Note: frame shift mutation from an ovarian tumor  
C:Genetics:  
A:Gene: GDB:BRCA1  
A:Cross-references: GDB:126611; OMIM:113705  
A:Map position: 17q21-17q21  
C:Superfamily: breast/ovarian cancer susceptibility protein BRCA1: RING finger homolo C:Keywords: mammary gland; nucleus; ovary; tumor suppressor; zinc finger  
F:20-70/Domain: RING finger homology <RNG>  
F:24-64/Region: zinc finger C3HC4 motif  
F:607-621/Region: nuclear location signal  
F:651-664/Region: nuclear location signal  
F:679-692/Region: nuclear location signal  
F:24,27,44,47/Binding site: zinc (Cys, His, Cys, Cys) #status predicted  
F:39,41,61,64/Binding site: zinc (Cys, His, Cys, Cys) #status predicted  
Query Match 99.9% ; Score 9642; DB 1; Length 1863;

	Best Local Similarity	99.9%	Pred. No. 0:	Matches 1862:	Conservative	0:	Mismatches	1:	Indels	0:	Gaps	0:
OY	1	MDLSALRVEEYONVIA	MOKILECPICELIK	EPVSTKCDHICRCKMLK	LNOKKGPSQ	60						
Db	1	MDLSALRVEEYONVIA	MOKILECPICELIK	EPVSTKCDHICRCKMLK	LNOKKGPSQ	60						
OY	61	CPCLKNDITKRSLO	ESTRESQVLEELLK	ICAFOLDTGLEANSYN	FAKKENNSPEHLK	120						
Db	61	CPCLKNDITKRSLO	ESTRESQVLEELLK	ICAFOLDTGLEANSYN	FAKKENNSPEHLK	120						
OY	121	EVSITIQSMGYRNR	ARKLLOSEPENPSLO	ETSLSVQJNSIGTV	TLTKQRIQOKSYI	180						
Db	121	EVSITIQSMGYRNR	ARKLLOSEPENPSLO	ETSLSVQJNSIGTV	TLTKQRIQOKSYI	180						
OY	181	ELGSDSEDTYKAT	VCYSGDOELLOIT	POGTDEISLDSAK	KAACEFSETDTNTEH	240						
Db	181	ELGSDSEDTYKAT	VCYSGDOELLOIT	POGTDEISLDSAK	KAACEFSETDTNTEH	240						
OY	241	PSNNDLNTTEKRA	ERHPEKYOSSVSNL	HVEPCGNTTHASS	LOHENSLLTKDRMNV	300						
Db	241	PSNNDLNTTEKRA	ERHPEKYOSSVSNL	HVEPCGNTTHASS	LOHENSLLTKDRMNV	300						
OY	301	KAEPFNKSKQPLA	SQHNRWAGSKETC	NDRTPTSEKAVD	LNADPLCEKRNKKO	LPC	360					
Db	301	KAEPFNKSKQPLA	SQHNRWAGSKETC	NDRTPTSEKAVD	LNADPLCEKRNKKO	LPC	360					
OY	361	SENPRDTEVPWIT	LNSIQKYNEMFSR	DELLGSDSDHGE	SESAKAVADVLVNE	420						
Db	361	SENPRDTEVPWIT	LNSIQKYNEMFSR	DELLGSDSDHGE	SESAKAVADVLVNE	420						
OY	421	EYSGSEKIDLLAS	PDHEALICKSERV	HSKSVESNIEDK	IFGKTYRKASLPM	SHVTEN	480					
Db	421	EYSGSEKIDLLAS	PDHEALICKSERV	HSKSVESNIEDK	IFGKTYRKASLPM	SHVTEN	480					
OY	481	LITGFVTEPQIIO	ERPLTNKLRKRRT	SGLHPDEFIKKAD	LAVOKTEMTINQ	NOTE	540					
Db	481	LITGFVTEPQIIO	ERPLTNKLRKRRT	SGLHPDEFIKKAD	LAVOKTEMTINQ	NOTE	540					
OY	541	ONGQVANTINS	GHEKTKGDSIQNE	KNPPIESLEKSA	FKTAEPISSISNM	LELNI	600					
Db	541	ONGQVANTINS	GHEKTKGDSIQNE	KNPPIESLEKSA	FKTAEPISSISNM	LELNI	600					
OY	601	HNSKAPKKNR	LRRKSTRIHAI	LELVSNLSPPN	CTELOIDSCSSSE	EIKKKRYNOMPY	660					
Db	601	HNSKAPKKNR	LRRKSTRIHAI	LELVSNLSPPN	CTELOIDSCSSSE	EIKKKRYNOMPY	660					
OY	661	RHSRLQIMECKE	EPATGAKSKNP	NEQTSKRHDS	DTFPELKLTNA	GSGTCKGNTSELKE	720					
Db	661	RHSRLQIMECKE	EPATGAKSKNP	NEQTSKRHDS	DTFPELKLTNA	GSGTCKGNTSELKE	720					
OY	721	FVNPSLPREEK	EKELETIVKSN	NAEDPKDMLSGE	EVLOTERSVSSIS	ISLVGTGYGTO	780					
Db	721	FVNPSLPREEK	EKELETIVKSN	NAEDPKDMLSGE	EVLOTERSVSSIS	ISLVGTGYGTO	780					
OY	781	ESISLLEVSTJG	KANTEPNKCVS	QCAAFENPKGL	IHGCSKDNKND	TEGEFRYPLGHEVNS	840					
Db	781	ESISLLEVSTJG	KANTEPNKCVS	QCAAFENPKGL	IHGCSKDNKND	TEGEFRYPLGHEVNS	840					
OY	841	RETSTEMEES	ELDAQYLONTF	KVSKROSALF	SNQNAEBCAT	FSAHSGSLAKQSPKYT	900					
Db	841	RETSTEMEES	ELDAQYLONTF	KVSKROSALF	SNQNAEBCAT	FSAHSGSLAKQSPKYT	900					
OY	901	FECDEKENE	OGKNEINIKPV	QTVNITAGPV	YGOKDPVDMN	KCSIKGSRCLSSQFNG	960					
Db	901	FECDEKENE	OGKNEINIKPV	QTVNITAGPV	YGOKDPVDMN	KCSIKGSRCLSSQFNG	960					
OY	961	NETGLITPNK	HGLLONPRI	PLPFIKSVK	TKCKKNLLEEN	FEHSHMSPEREMENIPI	1020					
Db	961	NETGLITPNK	HGLLONPRI	PLPFIKSVK	TKCKKNLLEEN	FEHSHMSPEREMENIPI	1020					
OY	1021	STVSTISNN	IRENFKEASS	NIINEVGS	STNEVGSS	INEIGSSDENIOAELGRN	GPKL	1080				
Db	1021	STVSTISNN	IRENFKEASS	NIINEVGS	STNEVGSS	INEIGSSDENIOAELGRN	GPKL	1080				

Db	1021	STVSTISNN	IRENFKEASS	NIINEVGSS	INEIGSSDENIOAELGRN	GPKL	1080			
OY	1081	NAMLRLGVLO	PEYVKQSLPG	SNCKHPEIKKOE	EEVYQTVNTDFSP	YLISDLLEQPMGSS	1140			
Db	1081	NAMLRLGVLO	PEYVKQSLPG	SNCKHPEIKKOE	EEVYQTVNTDFSP	YLISDLLEQPMGSS	1140			
OY	1141	HASQVSETP	DDLLDDGEIK	EDTSPAENDIK	ESSAVFSKYOKG	ELSRSPFTHTLAAQ	1200			
Db	1141	HASQVSETP	DDLLDDGEIK	EDTSPAENDIK	ESSAVFSKYOKG	ELSRSPFTHTLAAQ	1200			
OY	1201	GVRGAKKLE	SSPENLSSDE	DELPCFOHLL	PFKVVNNIP	QSRHSTVATECLSKNTEENL	1260			
Db	1201	GVRGAKKLE	SSPENLSSDE	DELPCFOHLL	PFKVVNNIP	QSRHSTVATECLSKNTEENL	1260			
OY	1261	LSLKNLND	CNSQVYI	LAKASOEHL	SEETKCSALF	SSQCSLEDTANTNTQDFLGS	1320			
Db	1261	LSLKNLND	CNSQVYI	LAKASOEHL	SEETKCSALF	SSQCSLEDTANTNTQDFLGS	1320			
OY	1321	SKOMRHQ	SESQGVGL	SDKELVSD	DEERG	TGLEENNOEQSMDSNLGEA	SGCESETSVSE	1380		
Db	1321	SKOMRHQ	SESQGVGL	SDKELVSD	DEERG	TGLEENNOEQSMDSNLGEA	SGCESETSVSE	1380		
OY	1381	DCSGLSQ	SDILITTO	QORDTMQNL	IKLOEMAE	LAVALLEOHESQPSNSYPS	ILISDSSALE	1440		
Db	1381	DCSGLSQ	SDILITTO	QORDTMQNL	IKLOEMAE	LAVALLEOHESQPSNSYPS	ILISDSSALE	1440		
OY	1441	DLRNPQ	STSEKAVL	TSOKSEYPI	SQNPESLADK	FEVSADSTSKNKEPC	VERESSPSK	1500		
Db	1441	DLRNPQ	STSEKAVL	TSOKSEYPI	SQNPESLADK	FEVSADSTSKNKEPC	VERESSPSK	1500		
OY	1501	CPSLDRMY	MHSCGSLQ	NRNYP	PSQELLKVV	VEEQULEESGPHDLTETSYL	PRODLEG	1560		
Db	1501	CPSLDRMY	MHSCGSLQ	NRNYP	PSQELLKVV	VEEQULEESGPHDLTETSYL	PRODLEG	1560		
OY	1561	TPYLES	GILFSD	PPSDSE	EDRAP	SAVGNIPSTSLKLPOLK	VAESQSPAAATTT	1620		
Db	1561	TPYLES	GILFSD	PPSDSE	EDRAP	SAVGNIPSTSLKLPOLK	VAESQSPAAATTT	1620		
OY	1621	DTAGYNAME	ESVSREK	PELTAST	ERYNKRMS	AVVSGLPPEEPM	LYKKPARKHITLTM	1680		
Db	1621	DTAGYNAME	ESVSREK	PELTAST	ERYNKRMS	AVVSGLPPEEPM	LYKKPARKHITLTM	1680		
OY	1681	TEETHV	VNKTDAE	FCERTLKY	FLGIAGK	VVVSYPVWTOS	IKERKMLNEHDE	EVRGDV	1740	
Db	1681	TEETHV	VNKTDAE	FCERTLKY	FLGIAGK	VVVSYPVWTOS	IKERKMLNEHDE	EVRGDV	1740	
OY	1741	VNCRNHQ	PKRARE	SQDRKIF	FGLETC	CGPPTN	PTDQLEWVQDLC	GA SVKELSSFTL	1800	
Db	1741	VNCRNHQ	PKRARE	SQDRKIF	FGLETC	CGPPTN	PTDQLEWVQDLC	GA SVKELSSFTL	1800	
OY	1801	GTCVHP	ITVVOP	DAMT	EDNGFHA	IGOMCEAP	VYTRREWYLD	SAVALYQCO	ELDTYLPOLIPH	1860
Db	1801	GTCVHP	ITVVOP	DAMT	EDNGFHA	IGOMCEAP	VYTRREWYLD	SAVALYQCO	ELDTYLPOLIPH	1860
OY	1861	SHY	1863							
Db	1861	SHY	1863							

RESULT 2

149350  
breast/ovarian cancer susceptibility protein BRCA1 - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 12-Nov-1999  
C:Accession: I49350  
R:Marquis, S.T.; Rajan, J.V.; Wynshaw-Boris, A.; Xu, J.; Yin, G.Y.; Abel, K.J.; Weber  
Nature Genet. 11, 17-26, 1995  
A:Title: The developmental pattern of Brcal expression implies a role in differential  
A:Reference number: I49350; WUID:96021028; PMID:7550308  
A:Accession: I49350  
A:Status: preliminary; translated from GR/EMBL/DBDUB  
A:Molecule type: mRNA  
A:Residues: 1-1812 <RES>

## RESULT 2

149350  
breast/ovarian cancer susceptibility protein BRCA1 - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 12-Nov-1999  
C:Accession: I49350  
R:Marquis, S.T.; Rajan, J.V.; Wynshaw-Boris, A.; Xu, J.; Yin, G.Y.; Abel, K.J.; Weber  
Nature Genet. 11, 17-26, 1995  
A:Title: The developmental pattern of Brcal expression implies a role in differential  
A:Reference number: I49350; MUID:96021028; PMID:7550308  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1812 <RES>

A:Cross-references: EMBL:U031625; NID:g988213; PIDN:AA017114.1; PID:g988214  
 C:Genetics:  
 A:Gene: Brcal  
 C:Superfamily: breast/ovarian cancer susceptibility protein BRCAL; RING finger homology  
 C:Keywords: zinc  
 F:20-70/Domain: RING finger homology <RNG>

Query Match 50.6%; Score 4882; DB 2; Length 1812;  
 Best local Similarity 56.1%; Pred. No 2.5e-209;  
 Matches 1046; Conservative 261; Mismatches 487; Indels 70; Gaps 32;

QY 1 MDLSALRVEEYQVYINAMOKILECPICLEDIKEPVSTKCDHIFCKCMKLNQKKGPSQ 60  
 Db 1 MDLSAVQIOEVQVNVLAHQKILECPICLEDIKEPVSTKCDHIFCKCMKLNQKKGPSQ 60  
 QY 61 CPLCKNDITKRSIQESTRFSQVLELLKICAFQDGTLEYANSYNAKKNENSPERLKD 120  
 Db 61 CPLCKNEITKRSIQESTRFSQVLELLKICAFQDGTLEYANSYNAKKNENSPERLKD 120  
 QY 121 EVSIIDSMGYRNRAKRLQSEPNPSLOETSLSVOLSLGTVRLTKRORIOPOKTSYI 180  
 Db 121 EASIIISVGRNVRRLPQVEPQGNATLKD-SLGVOLSLNLTIVSVKKNQTOPRKKSYYI 179  
 QY 181 ELGSDSEEDTVNKAIVCSVDQELLQITPOGTDEISLDSAKKACEFSETDTNTEHHQ 240  
 Db 180 ELDSDESSEETVTKPGDCSVRDELLQITAPQEADEGLKHSABACEFSE-GIRNIEHHQ 238  
 QY 241 PSMNDITMTEKRAAEHPEKYOSSVSNLHVEPCGNTTHASSIQHENSILLTKDMNVE 300  
 Db 239 CS-DLNPENHATERHREKQCSISINVCVEPCGDAAHSLQEPSSILLIEDRNAE 297  
 QY 301 KAFCKNKSQKOPGLARSOHNMASSKETCNDRTPTSTEEKYVDLADNAPLCEKKNKOKLPC 360  
 Db 298 KAFCKNKSQKOPGLARSOHNMASSKETCNDRTPTSTEEKYVDLADNAPLCEKKNKOKLPC 357  
 QY 361 SENRPTDEVPWITLNSIOLKVNEMFSRSDELLGSDSDHGESENAKVAADVLDVLENYD 420  
 Db 358 PENSGATTPDPWITLNSIOLKVNEMFSRSDELLGSDSDHGESENAKVAADVLDVLENYD 417  
 QY 421 EYSGSESEIDLLASDHEALICKSEVHKSVSNSIEDKTFGKTYRKASLPMLSHYEN 480  
 Db 418 GGFSSSRKTDVLPDHHHTLCKSGRDFSKPVEDNITSDKIFGKSYSORCKGRPLNHYTE- 476  
 QY 481 LITGAVTEPQIOTERPLTNLKRKRRTSGLHPEDFIKKADLA-VOKTPREMNOGNOT 539  
 Db 477 -IIGTITTEPQIOTERPLTNLKRKRRTSGLHPEDFIKKADLA-VOKTPREMNOGNOT 533  
 QY 540 EQNGQVYNTINSCHENKTKGDSIQNEKNPNPISLEKESAFKTKAPDISISSINMELELN 599  
 Db 534 EPNEQAVSTTSCQENKIAGSNIQKESAHPTESLKRREPASTAGAKSISNSVSDLEELN 593  
 QY 600 INHSAKPKNRRLKRSKSTRITIALELVNSRLSPMCTELQIDSCSSSEETKAKKNOMP 659  
 Db 594 VHSKSAKPKNRRLKRSKSTRITIALELVNSRLSPMCTELQIDSCSSSEETKAKKNOMP 652  
 QY 660 VRHSRLQIMEGKEPAGAKSNKPNQOTSRRSDSTPELTKLTNAPGSTCKSNTSELK 719  
 Db 653 AGHLRREPQIOTERPADAKK-NEPNEHTRKRASDAFPEEKLMNKAAGLLTSCSSPRKKSQ 711  
 QY 720 EYVNPSPREEKEKLETYKVSNNADPKDMLSGE-RVLQTERSVSSSISLVPGTDYQ 778  
 Db 712 GPVNPSPREEKEKLETYKVSNNADPKDMLSGE-RVLQTERSVSSSISLVPGTDYQ 770  
 QY 779 TQESITLLEVSTLGRKTEPNKVSQCAFENKGLIHGSKNRNDTEGFKPLGHEVN 838  
 Db 771 TQNSVSVLAHVTYRYARTGSAQCTQFVSENKELVHG-SNNAGSTGTELPKPLHAILN 829  
 QY 839 HSRETIEMESELDAQYLONTFKVSKROSFALFSPNGANECAATFSAHSGSLKQSPK 898  
 Db 830 LSGE-KYEMDESLDQYLONTFKVSKROSFALFSPNGANECAATFSAHSGSLKQSPK 894  
 QY 899 VTEFCQKEENQCKNESNITKPVQTVITAGPVGQKDPVDNAKCSIKGSRFCILSSQF 958  
 Db 899 VTEFCQKEENQCKNESNITKPVQTVITAGPVGQKDPVDNAKCSIKGSRFCILSSQF 958

Db 885 VTAKGKQK-ERQGEFEIISHVQAVATVGLPYPCQEGKLAADTMCD--RGCRILCPSSH 941  
 QY 959 RGNMTGLITPNKGLLQNPRIPLPIKFSVETKCKNLLEENFEHSMSPREKNGEN 1018  
 Db 942 RSGENGLSATGKGLISONSHPKQSVSIFRSITDNKPLTERFPHSTSTEAANGEN 1001  
 QY 1019 I-PSVSTISRNNIREVFEKASSNINEVSSSTNEVSSINEIGSSDEMIQAEIGNRG 1077  
 Db 1002 ILDSTHYTSLN-RGNACQEACS-----GSINEVCSGTGDSFPQOLGRNRG 1046  
 QY 1078 PKLNLMLRLGLVLOPEYVKOSLPGSCKNHPKIKQEEYEVQVYNTDFSPYLISDNEQPM 1137  
 Db 1047 PKVNTVPLDLSMOPGVQCOGVPSV-KYLEIKKQEGEAVC-----ADSPRLTFSDHLEQSM 1101  
 QY 1138 GSSHASQVCEFTPDDLLDGEIEKEDTSFANDIKESAVSKSVQGLSRSPTFHTH 1197  
 Db 1102 -SGKRVQVCEFTPDDLLDGEIEKEDTSFANDIKESAVSKSVQGLSRSPTFHTH 1160  
 QY 1198 LAQYRGRKARKLESSEENLSEDEELPCFQHLIFGKNNIPQSTRSHVATECLSKNTE 1257  
 Db 1161 KSGSLIRASKRLESSEEDSDTEDEDLPCFQHL-SRISNTP-ELTRCSSAVTORMPKAE 1218  
 QY 1258 ENILSLKNSLNDSCNOYILAKASQEHHLSEETKCSALFSSQCSLEDLTANTYQDPL 1317  
 Db 1219 GTQAPMKGSSSDCENNEVIMTEASQEHQFSDPCSGSMFSSQHSAAQGSTANANSQDSNF 1278  
 QY 1318 IGSSKQMRHSESOQVGLSDKELVSDDEERGDTGLEENNOEQSDMSNLGA-ASQCSSE 1376  
 Db 1279 IPPSKQSHQCGDEAFSLSKELISDNEEMATCLEEDNDEE--DSTIIPSEASYESET 1336  
 QY 1377 SVSEDSGLSSQSDILITTOORDTMQNLILKLOEAMELEAVLEBOHGSOPNSYISIDS 1436  
 Db 1337 NLSDED-----SQSDILITTOQRATMKYNLIKLOEAMLEAVLEBOHGSOPNSYISIDS 1392  
 QY 1437 SALEDLRNPEQSTSEKAVLTQSKSEYVPIQNPBGLSADKEVYS-ADSSSTSKNKEPQVER 1495  
 Db 1393 CALEDLRNPEQSTSEKAVLTQSKSEYVPIQNPBGLSADKEVYS-ADSSSTSKNKEPQVER 1452  
 QY 1496 SSPSKCPSLDDRYMYMSCSSLQNRNTPPSOEBELIKVYDVEOQLEBGPDLTETSLPR 1555  
 Db 1453 PSPFKSPPLAGRSASAHGCSNHLQKRNPSQDELLQPAQSE---ASSEPHNSTGQSLPR 1508  
 QY 1556 QDLGCTPYLESIGLISLPSD-PPESDPSSEDRAPESARVQNPISSTALKVPOLKVAESAOSP 1614  
 Db 1509 RELEGTPYLGSLISLSSRPSESEK-----EPAHIGTTPASSTALKIPGQVAFNSAA 1564  
 QY 1615 AAHTTDTAGYNAMEESVSRKPELTAISTERVYKRMNVVSGLTPEEFMLVYFARKNHT 1674  
 Db 1565 AGAD-----KAVGVIVSKIKPELTSSEERADRDISMVVSGLTPREVMYVQKFAERYL 1617  
 QY 1675 TLINLTETEETHVYVMTDAEFVCEERTIKYFLGIAGKVVSYTYVWQSIKERMLNEHD 1734  
 Db 1618 TLIDATITEETHVYVMTDAEFVCEERTIKYFLGIAGKVVSYTYVWQSIKERMLNEHD 1677  
 QY 1735 EYVRDGVNCRHNGHOPKRAKRSQDRKIFRGLEICCYGFTFMPDQLEEMVQLOLCASVYKE 1794  
 Db 1678 EYVGDVYVNGHNGHOPKRAKRSQDRK-KLFKGLQVYCCPEFTMPDELEFRLQLOLCASVYKE 1736  
 QY 1795 LSSFTLGTGVHPIVVQPDAMTEDENGFHAIGQCEAPVYTRREVVLDSVALYQCOELDTYL 1854  
 Db 1737 LPSTLHTDGTGHAIVVIVQPSAWTEDSNCPDQGLCKARLVMDVLDLSSTYRCRGIDAYL 1796  
 QY 1855 IPOI 1858  
 Db 1797 VQNI 1800

RESULT 3  
 T16870  
 hypothetical protein t13h2.3 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 02-Sep-2000  
 C:Accession: T16870

R.Wu, X.  
 submitted to the EMBL Data Library, October 1995  
 A:Description: The sequence of C. elegans cosmid T13H2.  
 A:Reference number: Z18593  
 A:Accession: Z185870  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Accession: 1-2447 <MW>  
 A:Cross-references: EMBL:U39653; NID:g1049397; PID:g1049400; PIDN:AA52494.1; GSPDB:GNOC  
 A:Experimental source: strain Bristol N2; clone T13H2  
 C:Genetics:  
 A:Gene: CESP-T13H2.3  
 A:Map position: X  
 A:Insertions: 18/3; 135/3; 176/3; 496/1; 693/1; 816/1; 1053/1; 1131/3; 1181/3; 1308/3; 1645  
 C:Superfamily: RING finger homology  
 F:158-207/Domain: RING finger homology <RRN>

Query Match 3.98; Score 373; DB 2; Length 2447;  
 Best Local Similarity 18.78; Pred. No. 1.5e-08;  
 Matches 371; Conservative 287; Mismatches 744; Indels 580; Gaps 95;

```

OY 2 DLSALRVEEVQNVINA-----MOKIL--ECPICLELKEPVST-KCHIFCKRC 47
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
127 EITLWADYVRKKNVLPKSCVCEADLVIGSFCDVCOELLQSGIMTKCGHRCDCQ 186
OY 48 MCLLNQKGPQCPCLKNDI-TRKSLQESTRFSQLELL-KIICAFQDGLLEYANS 104
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
187 ILVAF-MRSGNTPCFCRQNLGSKRELQDPRFDQILQYVESRSIVGRMMAREHEKD 244
OY 105 YNFAK-----ENNS-----PEHLKD-----EVSIGQMGTRNRAKRLQSE 141
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
245 YVEGKGYIEGSDMNKRYGIDNPCKLAPRPLKSAKRRKIRWFHESDEGQVRAKWESE 304
OY 142 PENPQLQESTLSVOLNLTGTVRLTKRQIOPKTSVYIELGSDSEDTYVNKATYCVG 201
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
305 KGAPREDQNTY-LENDKEST-SVAAEKYLEGEMPFLEIKSSDEQI-----DLDD 355
OY 202 QELLQITPGTRDEISLDS-----AKKAACEFSETDVTNTEHHQPSNND-----LN 247
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
356 EE-----ESMLDQFEISDNEDYKSPCSTKTKNRSRSDSSDNDSDNRNEIQ 404
OY 248 TTERKAAREHPRKYGSSVSNLHVEPCGNTNTHASSLQIHENSILLTKDRMNVKAEFCNK 307
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
405 KKKRMMKKNVPRKTOGSDVSN-----ESDEDAQSGEVATKLIKESKKRPGCR 452
OY 308 SKQ---PGLARQHNRMAGSKET-CNDRRTPTSEKKVLDLNP--LCEKREMNKOKLPSC 361
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
453 PKKKAPPELIE-----CDIPTPSDSLTSDEEDDNAADPYAFVFOKEFNRD--PRR 503
OY 362 ENPRDTEVP---WITLNSIQKYNEMFSRDEL-LGSDDS--HDG-----ES 403
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
504 DGHPRKDLKYNEDFMIDM---HQVDRKFEKDEGIHIVISDENSEHESEDAEDRESSIDS 560
OY 404 ESNNAVADVL---DVLNEVDEYSGSSSEKIDLLASDPHEALICKSEVHKSVSNEIDETI 460
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
561 EHEKTSIKFSLSRQPLPNPTSVDDCOQVITVVKDKQKQASATSKPGETSPDSSSKLEERP 620
OY 461 FGKTRKKAASLPNLSHVT-----ENLIGAFVTEP-----QIIQERPL 498
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
621 DKVSEVEDDEMPREHITADKGTDFLNNIMHEDEMTCGYLFRPDQITISRPKYORAGC 680
OY 499 TNNLKRRRPTSGLR-----PEDFIKADLAV-----OKTPEMINQNTQEQNQ 544
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
661 TNRLSMNVCPKAVLKEKKKLVIPPEDEYIISDEYTLSDSETSPSAAEQSGSESEAP 740
OY 545 VVNITNSHENKTKDSDSTONEKNPPIESLEKESAFKTAEDIS-----SISNM 594
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
741 STIKTSCTERTQSSSEPTSRDRKMHKRK-LDTFRRLKVLASDLSDEDFVSIDDN 799
OY 595 EL-----ELNTHNSAKPKNRLRKSSPTTHIALELVYSRNLSPNCTELODSCS--SS 647
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
800 ELVATGKPIIKHKVYDYSANRPKSN-----LDFTGRR-----NARETPMEISIRLAE 848

```

```

OY 648 EETKRRKNQMPRVHRSLNQLMEGKEPATGAKSNKPNEQTSKRH-DSDEPELKLINAP 706
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
849 EQVAHEBEK-----IHRROVLEAVEAA--SKLLNVYDTEEBEIEEBEETPEEVEKV- 901
OY 707 GSFTKCSNLSKEFENP--SLPRE-----KEEKLEYKVSNNAEEDKDLMSGERLYQ 759
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
902 ----ASPAPAIAT-ENPTTSTAPEEGVAMETPIE--EIFFDDEPCSSAQAAQRELI 953
OY 760 TERSVSSSISLVPGDYTOQSI--SLVEVSTLKAMT--EPNKCVCQCAFENPKGLI 815
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
954 IER-----VKEEQIILEDLSIQKRRKSSKTVESESEREKQEPRIEDENES 999
OY 816 HGCSKDRNDTEGFKYPLIGHEVNSHRETSIEMESELDAQYLONTFKVSKQSFALFSNP 875
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1000 EQOKKQADNPT--VEVDKESSEASSSPKSDPEDETLDAQ--SKTVKISLKHETV----- 1051
OY 876 GNAEECAFFSAHSG-----SLKKQSPKTFE-----CEQ 905
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1052 --SDEIEDFDTRFGFVATADAKMIKRTIGEVSTELKLVAQDPATDEVLAGFCVR 1109
OY 906 KEENQ-----GK-----NESNKPQVQVNTAGFPVYQOKDKPYD----- 940
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1110 NTDQESTIKETGKRTNKNRPDESMK-----NPRESP--AAKHPPVPRKLPNTNIEMK 1161
OY 941 ----NAKCSIKGSRFCLSQFRQNETGLTPNKHG-----LQ 975
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1162 IERAHVVKYKHVDMPELHKILIALQKQIALATCANLSQPVYTPREHBAQVOLLHNLQ 1221
OY 976 NPYRIPRPLPISFYKT--KCKKNLLEENFEHSHSPREM-GNINIFSYSTISRNIR 1032
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1222 NPSILRPLNPDQFALTTHKAOQA100QRAOKAQOTKELAAQAEQARVEELARRIA 1281
OY 1033 ENVFKAASSNINENQSTNEV-GSSINEIGSDENIOAELGRNRPKLTMLRLGLQ 1091
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1282 Q-----EDAKALRQGEQMSNVSGIPV-----SSDQMASSNAQOQG----- 1319
OY 1092 EYVKQSLPGSNCKHPEIKKQEEVEVQVYNTD--FSPYLSIDNLEQPMGSSHASQVSET 1149
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1320 ----LEENQTTTNSLSLTPRLTLDN--SHLGESQQLPVIESI 1357
OY 1150 PDLLDDGEIKEDTFAENDI-KESSAVFSKSVQKGLSRSPPTTHHLAQGRGAKK 1208
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1358 QGSTSE--ALKSEYWKDPIITLPASTVSKS--SAPATRRS-----RPPSSYDRSSP 1408
OY 1209 LESSEENLSDEDELPCOHLFGVNNIPNOSTHSTVAEELCKNTEEMLSLKSIN 1268
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1409 SVVIRERLGSQ-----ALINRPNRCNDKRSRSPISRRAVEIV-----RIN 1452
OY 1269 D-CSNOVILAKASQEHHLSEETKCSASLFPSSQSELEDLTANTPTQDFLIGSSKOMHQ 1327
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1453 DHQNETIL--AGNITHIVETT-----ILEGTS-----IGQDSTIRYD 1489
OY 1328 SESQGVGLSDKELVADDEBERGCLGENNOBEDSMQSNLGEAASGCESETSVSEDCGLSS 1387
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1490 GECSTQYITDKTIDLDSKNGTNVDE-----E 1516
OY 1388 QSDILTLTOORD-----TM-----QHNILKLOQEMA--ELEAVLEQ 1420
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1517 QSNVLAKRLBNDRNREMLYANKYHPSYMLAGNLSINRHN--KVQOVLASQELQDILAR 1574
OY 1421 HGS-OPSNSYPSIT-----SDSSALDELNRNPEOSTSEKAVLTSQSSSEYPISONEGL 1472
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1575 HTSGAVSQTYQVEVGGEGVACAGTSALGETDEDDOVEBEFPYDQELAKKILKQOGL 1634
OY 1473 SADKFEVNASDSTSKNKEPVERSSPK-----CSSLDRMYHSCSSQLQ 1518
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1635 ESESDSDDEDMVYDNOVGSVIARPAHKKRETRKKKNIFVNPPIKRIKKRYDK-KIEME 1693
OY 1519 NKNYPSOEELIKVYDVE-----EQOLESSGPHDLTETVYL--PRO----- 1556
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1694 RAKYIARIKSQMASIRIYAVPOKPLQOPATQOQVYRGPGKHSAAAAAARATPKPKAMSN 1753
OY 1557 ----DLEGTPLYLESGL-----SLFSD--DPESDP-----SEDRAPES--AR 1589

```

Db 1754 VSAPFTKTAVPNVAAGHONOOOLYSDMAAPQSTPIRTTPQPGTSGAPQAQTQSHLAQ 1813  
QY 1590 VGNIPSTSLAKVYPOLK-----VAESAASPAAHHTTDTGAYNAMEESVSRE 1635  
Db 1814 LGQVNVNANQOQAAPQOQGNTTAOLOLOMAQAANAVAOTAAQAQAAYAAE-AAVQAOQAQAARA 1872  
QY 1636 KP 1637  
Db 1873 AP 1874

RESULT 4  
T42205  
breast cancer susceptibility protein BRCA2 - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 21-Jan-2000  
C:Accession: T42205  
R:McAllister, K.A.; Haugen-Strano, A.; Hagevik, S.; Collins, N.K.; Brownlee, H.; Futreal,  
submitted to the EMBL data library, February 1997  
A:Description: Characterization of the mouse and rat homologs of the BRCA2 breast cancer  
A:Reference number: 222073  
A:Accession: T42205  
A:Status: preliminary; translated from GR/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-3329 <MCA>  
A:Cross-references: EMBL:089652; NID:92443438; PID:92443439; PIDN:AMB71377.1  
A:Experimental source: strain CD1; 129Sv; ICR Swiss  
C:Genetics:  
A:Gene: BRCA2  
C:Superfamily: breast cancer tumor suppressor BRCA2

Query Match 3.8%; Score 365.5; DB 2; Length 3329;  
Best Local Similarity 19.4%; Pred. No. 5,1e-08;  
Matches 402; Conservative 290; Mismatches 750; Indels 635; Gaps 93;

QY 72 SLESTFSPQLVEELKTIIC-----AFQDTGLEVANSYNFAKKENNS--PEHLKD 120  
Db 466 SLEDSIGKQMVSTSQACLSPISIRKSIKREPLDELTGTWSDGMTSTFEEHEAS 525  
QY 121 EVSTIISGMRYNRKAKRLLOSEPNS---LQESTLSVQLSNIGTVTLTKQIOPQKT 176  
Db 526 ACGIGITACQREDSICPSSVDGTGWPPTLTDTSATVK--NAGLSTLTKNKK-- 577  
QY 177 SVYIELGSDSEPTVKNATYCSVGDDELQITPQGTDEISLDSAKKAECSETDV--- 233  
Db 578 -----KFIY-SVSDASLQGGKKIQTHROLELTNL-SAOLEASFEVPLT 619  
QY 234 -TNTENHQPNSNDINTTEKRAERHPEKRYGSSVSNLHVEPCGTNTHASSLOHENSLS- 290  
Db 620 FTNNNSGIPDSSD---KKRCLPNDPE-----EPSLTNSFGTATSKETSIYIHA 663  
QY 291 LITFRNRNVKAEFCNKSQGLARSOHNKMAWAGSKETC-NDKRTPEP---EKYVDLNA-- 344  
Db 664 LISOD-LNDKEALYIEEKPOPYTAREADFLCLPERECENDQKSPKVSNGKEKVLASACL 722  
QY 345 -----DPLCEKKEWNKQKL-PCSENPRTEDVDPV----- 372  
Db 723 PSAVQLSSISFEQENLGDHNGSTLKLTPPSKRLPSTKADWVSREKMKMPEKLQCESC 782  
QY 373 ---ITLNSSTQKVNEMFSRSDDELGSDSDHGESESNAKV-----ADVLVDLNEVDY 422  
Db 783 KVNTELSKNLVEVEILISE-----NSKTPGLLPGENITIEVASKRSQ 827  
QY 423 SSGSEKIDLLASDPHEALICKSER---VHKSVESNIEDIK-----FKTYRKAKSLP 472  
Db 828 FNONAKI-----VIOKDOGSPFISEVAVNMNSEEPLPDGNNFAPQVTNKNKP 877  
QY 473 NLSHTENLIGAFVTEPQIIOERPL--TNKLKRRKRRPTGLHPEDEIKKADLAVQKTP 530  
Db 878 DLGSSVE-----LQEDLSHTGQPSLKNSPMAVDEVDNAHAQAQVILITDSD 924  
QY 531 MINQGTQTEONGOVNNTNSGHENKT-KGDSIONEKNPNPDESLEKESAFKTAEPRISS 589

Db 925 SLAVVHDYTEKSRNNIRHOAKGTEDKDFKSSVLSNMKSDGNSDCKSEF---LDVPLN 981  
QY 590 -----SISNMELEINHNSKAPKK--NLRKRSSR-----HIALELVSRNLSP 634  
Db 982 HNFQGSFRYASNKEIKLSEHNVAKSKMFEKIEQYPRLLACIDIVNTLPLANOKKISEP 1041  
QY 635 NCTELQIDSCSSSEIEIKKKYQNPVRRSRNLQMEGEPATGAKKSNKPNEQTSKRDS 694  
Db 1042 HIFDLK-----SVTTVSQSHNQSVSH-----EDT 1067  
QY 695 DTFPELKITNAPGSFTQCSNTSELKEFYVNSLPREEKEKLETKY-----VSNMADP 747  
Db 1068 DTAPQM-LSSKODPHSNLTTSOAKETTELSTLIEESGQPEFQFRKPSHIAQNTSEVP 1126  
QY 748 KDLMLSGEVRVQY-ERSVSESSISLVQGTQVQTESISLLEVSTLIGAKTEPNKCVSQCA 806  
Db 1127 GNOVY-----VISTASKKEDTDLHPVDPSPVGQTDHDKQFEGSA-GYKQSPHILLEDTCN 1181  
QY 807 AFENPKGLIHQCSKDNNDTE--GFKYPLGHEVHNSRETSLIEMESELDAQYLQNTFKVS 864  
Db 1182 --KMT-----SCFLPNINEMFEGFCALGTKLSVSN-----A 1213  
QY 865 KROSFALFSPDGNAAEEBCA-----TFSAHSGS-----LKQSPKVTVE---CEQK 906  
Db 1214 LKRAKMLFSDIENSEEPSAKVGPFGFSSAHHDVASVFIKQNTKESPFDEKSSKQVPT 1273  
QY 907 EENO-----CKN-ESNKKPYQ-----TVNITAGPPV---QOK 935  
Db 1274 LQNNIEMTTCIFVGRNEPEYKIKNTKHEDSYTSQRRNLNENSDGMSSTSG-PVYIHKGDS 1332  
QY 936 DKPVD-NAKCSIKGSRFLCSSQFRGNETGLITPNKKGHLQNPRIPLPEIKSFVTKC 994  
Db 1333 DLPAQDQSK-----PESC--TOYAREENOQIKENISDICTLETKMAEFTCMKSDQKQL 1385  
QY 995 KKNLEENFEHBHNSPERENGENIPSTVSTISNNIRENVYKFASSNNINEVGSSTNEV 1054  
Db 1386 PSDMEONIKREFNIS-----FOTASGKNTR--VSKESLKNKSVNIFNREIDE- 1429  
QY 1055 GSSINELGSSDENQALGRNGRPLAMLRIGYLOPEYVYKQSLGSGCKRPELK-QQY 1113  
Db 1430 ---LTVISDS-----LNSKILHGNKKQMH-----TSCHRKAISIKKVF 1465  
QY 1114 EEVQVQVNTDFSP-----YLISDNLQOPMSSHAS-----QVCSPTPD---DLIDQGE- 1158  
Db 1466 EDHFPYIVSQLPQCHPVEYIEISTKEPTLISFTAGSKYKIMQESLDKKNLFDETQY 1525  
QY 1159 IKEDTFAEND--IKESSAVFSKSVQKGLSRSPSPFTTHLAQYRGAKKLESSENL 1216  
Db 1526 VRKTASFSGSKPLKDSKKELTLAYEKIEVY-----ASKCEMGNFV 1567  
QY 1217 SSEDDELPCFOHLLFGKVNINIPQSTRHSTVATE-----CLSKN---TEENL 1260  
Db 1568 SKETEMLPQOYNNHYRQTEMILKTSNGTSKVOENIENNVEKNPRICICQSSYPVTVDISA 1627  
QY 1261 LS--LKNSLDSCNOVLAK-----ASQEH- 1284  
Db 1628 LAYTEBSRKTQVARESSLGGRKWLREOGDKLGRNRTIKIECVKEHEDFAGNASTYHSL 1687  
QY 1285 -----LSEETKCSASLFPSSQSELDL----- 1306  
Db 1688 VIIRTEIDTNHVSQNVSTLSDPNVCHSYLSQSSPFCCHDMHNDSGYFLKNKIDSDVP 1747  
QY 1307 -----TAN-----TNTQ-----DPLIGSSK-- 1322  
Db 1748 DMNRAGNTTISPRYSATKERNLHPQTJNEYCVQKLENTNTPHANKDVAIDPFLDSSNCK 1807  
QY 1323 -----QMRHQSQGVGLSDKELVSDDEERGTLGLENNOEODSMDSNLGEAAGCSEF 1376  
Db 1808 VGSIVFTTAHQE-----RTKEIVDNCYKIY-----EQNRQSKPRPTCQTSCHKVL 1855  
QY 1377 SVEDCGSSQSDILTTQQRDTM-OHNLIKLOQEMAELEAVLEQHGQSPNSYPTSIISD 1435

Db 1856 DSKDFICPSSGDCVINSKRSFCPHNEIIOHNS-----MFLKKAATPPVGLT 1908  
 Qy 1436 SSALEDLRNDEQ-----STSEKAVLTSOKSSEYPIGONPEGLSADKFEVSAD 1483  
 Db 1909 WPTSKSIRERPOAHNSRTYIGFSTASGKAIQYSDASLE-KARQVSEMDGAKOLSSMV 1967  
 Qy 1484 STSKNEP--GVERSSPKCPSLDDRYMHSKCSGLQ-----NRNP----- 1523  
 Db 1968 SLEGNEKPHHSVAKRENS-----VYHSTQGVLSLPKPLPGNVNVSVEFGSTAGK 2017  
 Qy 1524 -----SEBELKYVDV-----EEOQLBESG-PHDLTETSYLPRLDLETPY-LBEGIS 1569  
 Db 2018 LVTVSESALHKKYKGMEEFDLIRTEHTLQHSPIPEDVSKILPQPCAIRTEPEYVNSKLQ 2077  
 Qy 1570 LPSDDPEPSDEPDEARAVGNIPSTSAKVLPOLKVAESAQSPAATHTTDAG----- 1624  
 Db 2078 KTYNDKSSLPNNK--ESSGSGTQSTIEVSLQLSQMRNODTQVLCTKKVSHKANLIGK 2135  
 Qy 1625 YNAMESSVAREKPEL-TASTERVNRKSMNVSGLTPEEFMLVYKFAKHHITLTNLITEE 1683  
 Db 2136 EQLPQNIKVKTDKMTFSDVPYKTVNGEYKSESENYFETEAVESAKAMEDELDLDS 2195  
 Qy 1684 THVYVKTAEPCERTLKYFLGIAGKVVSTFWYQSIKERKMLNHD--FEVRGDV 1741  
 Db 2196 QTHAKCSL--FTCPQNETLFSNRTKRGVTVDAVGQPPKRSLLNEFDRIIESKGS 2252  
 Qy 1742 NGRNHGPKRARES--ODRKIF--RGLEICCYGP 1772  
 Db 2253 T-----PSKSTPDGTIVKDRSLFTHHMSLEPVTCGP 2283

RESULT 5  
 T42207  
 breast cancer susceptibility protein BRCA2 - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 03-Dec-1999 #sequence: revision 03-Dec-1999 #text: change 21-Jan-2000  
 C:Accession: T42207  
 R:McAllister, K.A.; Haugen-Strano, A.; Hagevik, S.; Collins, N.K.; Brownlee, H.; Futreal, S.  
 Submitted to the EMBL Data Library, February 1997  
 A:Description: Characterization of the mouse and rat homologs of the BRCA2 breast cancer  
 A:Reference number: 222073  
 A:Accession: T42207  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-3343 <MCA>  
 A:Cross-references: EMBL:U89653; NID:q2443440; PID:q2443441; PIDN:AAB71378.1  
 A:Experimental source: strain Sprague-Dawley; testes  
 C:Genetics:  
 A:Gene: BRCA2  
 C:Superfamily: breast cancer tumor suppressor BRCA2

Query Match 3.7%; Score 361.5; DB 2; Length 3343;  
 Best local similarity 19.2%; Pred. No. 7.7e-08;  
 Matches 379; Conservative 297; Mismatches 715; Indels 579; Gaps 96;  
 Qy 66 NDITRSLOESTRFQSLVEELKIIICAFOLDLTGLVANSYNAFKKENSSEHL-KDEVST 124  
 Db 1087 NHILTPSKAEITELSTILE-----ESGSGEFTQFNK--PSHIAQNNTSA 1129  
 Qy 125 IOSMGRNRAKRLQSEPPNSLOETSLVSQLSNLGTVRLTKRQROPKTSVYLEGS 184  
 Db 1130 V--LGMQMAVVRASEEMKDVLDL--LPLNPSSVQIDHNKKFEELVYGQS-----SS 1179  
 Qy 185 DSSEPTVNRKATYCSVQDQELL--QITPOGTRDEISLDSAKKAACESETDVNTE----- 237  
 Db 1180 HLEPTDGNONTSCFLPIKEMEBGFGSALGTRKLSVNEALRKAMKLPSEDIENISEPSTK 1239  
 Qy 238 -----HH-----OPSNDLNTTEKRA-----AERPEKY 261  
 Db 1240 VGRPGSSCAHHDSVAVFKIKQNTDKSPDEKSSCQVTVQNNKEMTTCILVDENPENY 1299  
 Qy 262 QGSSVSNLHVEPCGTTTHASSLQHESSLLLTDRANVKEAEFCNK--SKQPLASQHN 319

Db 1300 -----VKNIKODNNVTGSQRNAYKLENSDV-----SKSSTGTVY INKGSDLP-FAAEKGN 1350  
 Qy 320 RMAGSKETC-----NDGRTPSTEEKVDL-----NAPLCEBREMKNOKIJCSENRDRED 369  
 Db 1351 KY--PESCTQYREENAQIKESVSDLTCLVYKAEETQHMSSDKEQLPSKMBQNNKE 1407  
 Qy 370 VPMILTNSSLOKVNEMFSRSDLELIGSDSDHSESENAKAVADLVYLVNEVDESGSSEKI 429  
 Db 1408 -----FNISFQ-----TASGKNIRVSKESLKNKSVILDO-QETELVYTSLSL 1448  
 Qy 430 DLLASDPHEALLKSER-----VHSKSVESENEDKIFGKTY--RKKASLPULSHVENTL 481  
 Db 1449 -----NSKILCIGINKMKHLSCHKKSI--NKK-KVEFEHPPIGVSLPQLOOYPEYE 1498  
 Qy 482 ILGAFYEP-----QIOE-----RPLNKLKRRRPPSLGP-----E 515  
 Db 1499 I--ESTKEPPLLSFHTASGKRYAKIMQESLDKKNLFDQYVYRKTTNGHOESKPLKORE 1556  
 Qy 516 DF-----IKKADLAOKTPREMINQTNQTE-----QNG-----QVANNITNSGHEKRTKGD 560  
 Db 1557 DYKERLLTAEKTELEVYASKCEEMQNFVSKQTEMLPQONDHMYRQTEENLTSNGSPKHYG- 1615  
 Qy 561 SION--EKNDPIESLEKDSAEKTKAEPIS-----SSISNNELELNHNSKA 605  
 Db 1616 NIENKIEKNR-ICCIQSSYFVTEEDSALACYGDSRKTCYGESSLSKGMKMLRQSDKL 1674  
 Qy 606 PKKNRLRKSSYTH-----IHAEIVVSRNLSPNCTELQIDSCSSSEIKKKKYNQ 657  
 Db 1675 GTRNTIETIQVKHTEDFQGNALYEHSLVYIRT-----ELDSHVEQNSTLESD 1725  
 Qy 658 MPVRSNRLOLMGKBPATGAKKSNKPNQTSKRHSDTFPELKLTMAGSFTKCSNTSE 717  
 Db 1726 PNVCHS-----YLSHSFCHDDMHNDGXYFLKDKIDSDVQPMKNTREGNAIPKISATKE 1781  
 Qy 718 LKFEVNSLPREKREKLETVKASNNABDPKDLMSGEVLOTBESVSSISLVPGRDY 777  
 Db 1782 IK--LHQYVNECVQKLET--NASPYAKNNLADISAMDLDRCKRGSVFT--TTH 1832  
 Qy 778 GTOESISLLEVSTLGRK-----TEPKVCVSOQ-AAEENPKGLIHGCSKN--RNDTE 827  
 Db 1833 -SOETVRMKEIFITDNCISKIVQNRRESKPTQCQTSCKALDNSEDFICPSSGSDVCINSPI 1891  
 Qy 828 GFRYPLGHEY-NHSRETS-----IMESELDQOYLQNTFKVSKRQSFALFS----- 873  
 Db 1892 AIFYPOSEOILQHNQSVGLKKAATPPVSLLEWMDCKSTRGSPQEHVPSRTYGFSTASG 1951  
 Qy 874 -----NPGNAEEFCATFSAHSGSLKQSPKVFEECQKRENGKKNENI-----K 918  
 Db 1952 KAVQVSDASLEKARQVSEIDGAKOLASMVSLGNEKSHSVKRESSVYHNTGVLSLR 2011  
 Qy 919 PVQTVNIT--AGFPVVGQDKPV-DNAKCSIKGSRFCLSSQFRGNETGLITPNKHGL 973  
 Db 2012 KTLRGVSSVFPVSGFTAGKLVTVSESLAHVKG-----MLEERD-----LITTEHTL 2060  
 Qy 974 LQNP-----YRIPLEPIKSFVTKCKKNLLENEEHSMSPEREMENIIPSTVSTIR 1028  
 Db 2061 QHSPTPEVDSKIP-----QCLESRTPEXSVS-----SKLOKTYNDKSR 2100  
 Qy 1029 NNIRENVKEAASSNINEVGSSTN-----EVGSINIEIGSSDENTQIADLGRNRGLNML 1084  
 Db 2101 -----SPSNKESGSSGNTQSLVSPQSLSEMERKQET--OSVLGTAVSQKTNIL 2148  
 Qy 1085 RLGVLOPEVYKQSLPSSNCKHPEIKQOEYEAVQVNVNDFSPYLLISDMLQDPMGSHASQ 1144  
 Db 2149 E-----KKQNL-P-QNKIKESNMKEPTSDVSMKTN-----VGE 2179  
 Qy 1145 VCSPTDILLDGEIKEDTSFAEND-IKESSAVFSK-----SVQKGLSRSPSPFTHTL 1198  
 Db 2180 YSKPEPENYFEIYAVIAKAFMEDELDTSEQTHAKCSLFAQPOREALINS-----PTRK 2234  
 Qy 1199 AGCYRRGAKKLSSSENLSSDEDELPCFOHLLFGKVNIN-PSQSTRHSTVATECLSKYTE 1257  
 Db 2235 RGMAGVAVAGOPPIKRSLLNE-----FDRIIESKGSILTYPKSTPDGTIKDRRLF----- 2284



QY 1258 ENLLSLKNSLNCDSNOVILAKASOEHLIS- EETKCSALFSSOCSELEDLTANTMTQDPF 1316  
 Db 2285 -----THHMSLEPYTCGPF-----CSSME-----ROETOSPH 2311  
 QY 1317 LIGSSKOMR---HOSSESQVGLSD-----KEIVSDEER---GTGLEENNQ 1356  
 Db 2312 VTSPAGOLQSKERHRSRAVGAOKSSNPTVSALNSEFTRHSHVSDKSTKVFPPEKVRFRH 2371  
 QY 1357 EEOGSMOS-NGEFAASGCESTVSED-----CSGLSSQSDIILTQOORIMOH 1402  
 Db 2372 RDEFPDKSNVNLBEGKNOKSADGVSEDOGNDSPQFNKDLMSLQANADLQDIRKKNERH 2431  
 QY 1403 NLIKLOEAMLEBAVLFEHOSQSPNSYPTIISDSALEDLRNEQO---STSEKAVLTQ 1458  
 Db 2432 HLCQPOESL-----YLTKSSTLPRISLOQAVGDS--VPSACSPKQLYMTGVSACISVNS 2484  
 QY 1459 KSESEY---PISON--PEGISADKFEVSADSS---TSKNKEPGEVRSRSPKC--PSLDD- 1506  
 Db 2485 KMEYFOFAIEDHFGKCALCAGKGFRLADGGMILPSDDGKAGKEEFYRALCDTPGVDPKL 2544  
 QY 1507 -----RWYHSCSG-----SLONRNPISOELIKV---VDEEQOLEESGPHD 1546  
 Db 2545 ISSVWVSNHTRMVLWKLAEMFAPPEFANRCLNPEKVLQLYRYDVEIDNSSRSALKK 2604  
 QY 1547 LTE-----TSYLPRODLEGTPEYLESGLSLEFSDDESDPEDEPESARVGNL----- 1593  
 Db 2605 ILERDPTAAKTLYLCVSDI-----ISLSTNVSTSGSKASSSEDSKNVDITELTDGW 2655  
 QY 1594 -----PSSTSALKVP---OLKVAESAQSP 1614  
 Db 2656 YAVKAQIDPLPLALVKSGLTVGOKIITOGAEIVGSDACAPLEAPDSLRLKLTISANSTRP 2715  
 QY 1615 AAMHT-----TDTAGYNAMESESREK---ELTASTERV--NKR 1649  
 Db 2716 ARHNSKIGFHDPRPPLPLSSLEFSDGNGVCDVLYQRYPLQWVEKTVSGSYIFERNR 2775  
 QY 1650 MSWVSGLPPEEFMLVYKFAKHHITLNLITEETHVAKTDAEFYCER 1699  
 Db 2776 -----EEBEKALRFPAEQOKKLEALFTK--VHTELKEHEEDLAQR 2813  
 RESULT 6  
 T13564  
 microtubule-associated protein homolog - fruit fly (Drosophila melanogaster)  
 N:Alternate names: hypothetical protein EG:49E4.1  
 C:Species: Drosophila melanogaster  
 C>Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 17-Nov-2000  
 C:Accession: T13564  
 R:Spanos, L.; Papadogiannakis, G.; Siden-Kiamos, I.; Louis, C.  
 submitted to the EMBL Data Library, April 1999  
 A:Description: Sequencing the distal X chromosome of Drosophila melanogaster.  
 A:Reference number: Z17689  
 A:Accession: T13564  
 A:Status: preliminary; translated from GB/EMBL/DDBB  
 A:Molecule type: DNA  
 A:Residues: 1-5327 <SPA>  
 A:Cross-references: EMBL:AL031128; PTDN:CAA20006.1  
 C:Genetics:  
 A:Cross-references: FlyBase:FBgn0025392  
 A:Intons: 24/2; 52/3; 104/3; 232/1; 1669/3; 2566/1; 4798/3; 5272/1  
 A:Note: EG:49E4.1  
 C:Superfamily: Drosophila 576k microtubule-associated protein homolog  
 Query Match 3.7%; Score 358.5; DB 2; Length 5327;  
 Best Local Similarity 17.9%; Pred. No. 2e-07;  
 Matches 362; Conservative 308; Mismatches 822; Indels 525; Gaps 75;  
 QY 31 IKPVPSTKCDHIFCKFCMLKLNOKK-----GPSQCLPKNDITRSLQES-TRF 79  
 Db 2355 VAEIVSSPIEDAEIMFESKIEIVKSSIALSLQSGSGKLOTDSPPVDVAGDSHAVASY 2414  
 QY 80 SOLVEELLKTIICAFQDLDGTGLEVANSYNFAKKNNSPEHLKDEVSITQSMGRRNRRKRLLO 139

Db 2415 STVTPTLTKPAELAQIQAATVSSPLDEALRTPSAPHI-----SRADSPAE 2461  
 QY 140 SEPPENSLQETSLSYQSLNIGTQVTLTKQRIQOKTSYVILEGSDSEEDTVKATTCV 199  
 Db 2462 CASEELIASQKSPQVLYKLESSRPAMVAESKDDAAQKSSVE-DLSPVASTEISRP--ASA 2518  
 QY 200 GDOELLQITPOGTRDEISLSAKKACEF-----SETDVTNTEHHOP---SNN 244  
 Db 2519 GETAASSPI-EEARKDFAEFQDAEKAVPLITELKGNLPTLSSPVDVAHAS-VQPAELSKV 2576  
 QY 245 DLNTEKRAERHPEKRYGSSVSNLHVEPCGTNTHASSLOHENSLLLTRDNVKEAEF 304  
 Db 2577 DIEKTASSPIDEAPKSLISGPAEREPESPASAKDAE-----SVEKSK- 2620  
 QY 305 CKNKSGKGLARSOHNRNRAKSKETCNDRRTSTEEKVVLNADPLCEKREMKOKLPCSENP 364  
 Db 2621 -DASRPVSVEST-----KADSTKGDISPSPEVLEGPDKDVEKSESSR--PPSVSA 2670  
 QY 365 RDTEDVPMITLNSSIQKVNEMFSKSDPLGSDSHODESESNKAVADVLDVNEVDEYSG 424  
 Db 2671 SITGD-----STKDVSPASVSEVKDEBHKAESRESSTAKVESVIDEAKSDBKSS 2722  
 QY 425 SSEKIDLASDPHEALICKSERVHKSVSSENIEDKIFGKTYRKKASLPNLSHVTENLIIG 484  
 Db 2723 SODS-----OKDEKSTLASKASRESVESSEKDDA-----EKSESRP----- 2760  
 QY 485 APTPEOPIIDERPILTNLKKRRPTSGLHPEDPLKAKDLAVOKTPENINGTNOTEQNG 544  
 Db 2761 -----ESVIASGEVPPRESK-----SPLDSKDTSRPSVSESYAE-----DEKSEQSR 2805  
 QY 545 VMNITNGHGNKTKGDSIQNEKNPNPLESEKESAFKTKAEPISSISNMELINHSK 604  
 Db 2806 RESVAESVKAADTKKDGSOASRPSVDELKDD--DEKESRQSTIGSHKAMSTMGDE 2863  
 QY 605 APRKNRLRRKSSRTHIALELVYSRNL---SPNCTELOIDSCSS--BEIKKKKNOM 658  
 Db 2864 SPMDKADKSKESPRESVASIKHEMTKDESPIGSR--DSVASIKSDITKGEKSP 2920  
 QY 659 PVHSHRLQIMECK---EPATGAKSKNPNQOTSKRIDSTFELKLTNAPGFTKCSNT 715  
 Db 2921 PSKEVSRPESVVSIIKDEKASRESVAESVKESSEKADATSAAPPKSHSR----- 2971  
 QY 716 SELKEFVNPSPREEKEKLETVKVSNNADPDLMLGCRVLOTESVSSSISLVPQT 775  
 Db 2972 -----ESVLSGLKDEGDKTTSRVSVADSTIDEKSLVSOE-ASRPSEASLSKDAAPSQ 3026  
 QY 776 DYGTQESISLLEVSTLGAKEPTEPNKCVSOCAEENPKGLIHGCSKDNRNDEGFKYPLGH 835  
 Db 3027 ETSRPESEVT--ESVKQOKSPVASKESRPAVSAN-----AKGSADSS--KQGRPE 3073  
 QY 836 EVNHSRSTSIEMESSELDQO-YLQNTFKVKRQSF-----LPSNPGMAE----- 880  
 Db 3074 SLPOKASGSIKDKSPLASKDEAKSEESRRESVAEOPLVSKESVSRPASVAESVDEA 3133  
 QY 881 -----CATFSAHSGSLKQSPVTFRECEBOKEENOCKNEINIPVOTVNTAG 928  
 Db 3134 EKSKESPLMSKASRAPASVAGSVKDEAK-----SKESRRESVAEKSPLSKESRSP 3187  
 QY 929 FPVVGQDKPVDNKAACSIR---GGSRPCLSSQFNGNETGLITNKGILQNPRIPLFP 985  
 Db 3188 ASVAESVKEADKSKESRRESGAESPLASKESRAPAS----- 3226  
 QY 986 IKSFVTKCKKNLLE---ENFEHSMSPREMGNGENIPSTVYITSRNIRENWFKEASS 1042  
 Db 3227 VASISIDKAEKSKESRRESVAESKSPLSKESR--PISVA---KSVDEAKESKEES 3279  
 QY 1043 NINEVGSSTNEVGSSTINEIGSSDENIOAELGRNRGKPLNMLGLVLO--PEYKO-SLP 1099  
 Db 3280 SRDSVAEKSLAKESKRAPASVAESVQDEAKSK---EESRRESVAEKSPLAYKESRSP 3335  
 QY 1100 GSNCK--HPEIKKQOEYEEVQVNTDFSPYL-----ISDNLEQ----- 1135

```

Db 3336 ASVAESIKDEAKESKESRESRY-AEKSPPLASKASRPTSVASVYKDEAKESKESRSRDS 3394
OY 1136 -----PMGSSHASOYCESRPDDLDGCEIKEDTSPFANIKIKESSAVESKSVQKELSRSP 1190
Db 3395 VAEKSPPLASKESKESRPAS-VAESYQDEAKESKESRES-VAESPLASK-----EASRPA 3447
OY 1191 SPFTTHLAQYRGARKLESSEN-----LSSEDEELPC-FOHLLFGKVNINPS 1239
Db 3448 S-----VAESYKDAEK--SKEESRESVAEKSPPLASKASRPASVAESVYKDEAKESK 3499
OY 1240 OSTRSTVATECLSKNTEENLLSKNSLNDSCNQVLLAKASOPHHLSEETKCSASFSSQ 1299
Db 3500 ESRESVAEKSPPLSKASRPTSVASVYD-----EAKSKESRESRESVAEKSSLASKE 3553
OY 1300 CSELDLANTNTODPFLIGSSKOMHOSQGVGLSDE----- 1339
Db 3554 ASRPASVAESVYDEAE--KSKESRESVAEKSPPLASKASRPASVAESVYKDEAKESK 3610
OY 1340 -----LVSDERGTGLENNQEQSMDNLGEAAGCESETSVSEDCGLSS 1387
Db 3611 VSRRESVAEKSPPLSKASRPTSVASVYKDEADKSKESRESREGCAKSPPLASMEASRPTS 3670
OY 1388 OSDILITQORDTQMHNHLKQOBMAELVLEQHSQPSNSTPSTISDSALEDLRNPQ 1447
Db 3671 VAESYKDETEKESKES--RESVTEKSPPLSKASRPT-SVAESYKDEAKESKESRRE 3726
OY 1448 STSEKAVILTOKSS-----EYPISONPGLSADKEV----- 1479
Db 3727 SVAEKSPPLASKESRPASVAESIKDEAECKOESRESMESKASISIGDOSLSKSET 3786
OY 1480 -----SADSTSKNKEPVERSSPKCP-----SLDDRVMHSCSGSLONRVPQ 1525
Db 3787 SRPDSVSVKQETEKPEGSATIDKQSVASRPESVAASAKDEKSPHSRESVADKSPDAS 3846
OY 1536 BELIKVYDEEQLE--ESGPHDLETSTLPRODLEG-----TPYESGISLFSDD----- 1574
Db 3847 KEASRSLVAETASAPIEGSPSIDLS-LP-LNLGEAKGKLPSTSPIDVAEGSFLEY 3904
OY 1575 -----PE-----SDPSDRAPESARVGNIPSSALSALVPOLKVAESQSPAATTOTAG 1625
Db 3905 KAESSPRPRAVLSPKPAEFSPDGTGHASTYVDEASPLLEIEVYO-----HTTGCVGA 3957
OY 1626 NA-----MEESVSR 1634
Db 3958 TGATAETDLDLTETKSETVTKOSETTLPETLTSKVESKVEVLESSVKQVEEKVQTSVQ 4017
OY 1635 EKPELTASTERVKNKMSMYVSGLTPEFPL-----Y 1666
Db 4018 AETVTYDLSLEOLTKKSS--EQLTEIKSVLDTNISNVTNLFSTAVETIEKKYQDVTEKYI 4074
OY 1667 KFARKH--HITLNLITEET-----THVVKTDAPF--VCERTLKY 1703
Db 4075 EKATHEVSEHYTTTSESTETQEKSSLDLGFSELRETHITTVGSPETVITICEDER- 4133
OY 1704 FLGIAGKRWVVSYFWYTOSIKERKMLNHDHFEVRCGV 1740
Db 4134 -----VLHDKREED--EHRFSPSPDV 4153

```

## RESULT 7

T30904

breast cancer tumor suppressor Brca2 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 21-Jan-2000

C:Accession: T30904

R:Conor, F.; Smith, A.; Wooster, R.; Stratton, M.; Dixon, A.; Campbell, E.; Tait, T.M.;

Hum. Mol. Genet. 6, 291-300, 1997

A:Title: Cloning, chromosomal mapping and expression pattern of the mouse Brca2 gene.

A:Reference number: 220931; MUID:97217789; PMID:9063750

C:Accession: T30904

A:Status: Preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-3329 &lt;CON&gt;

A:Cross-references: EMBL:U82270; NID:g1854950; PID:g1854951; PIDN:AA848306.1  
C:Genetics:  
A:Gene: Brca2  
A:Map position: 5  
C:Superfamily: Breast cancer tumor suppressor BRCA2  
C:Keywords: tumor suppressor

Query Match 3.7%; Score 357.5; DB 2: Length 3329;  
Best Local Similarity 19.2%; Pred. No. 1.2e-07;  
Matches 401; Conservative 292; Mismatches 742; Indels 649; Gaps 94;

```

OY 72 SLQESTRFSQLVEELIKITC-----AFQDTGLEVANSYNFAKKENNS--PEHLKD 120
Db 466 SLEDSTIAGQWVSRTSOACLSPTIRKSTFFKMRPELDLETIGVFSDSMTNSTFEHEAS 525
OY 121 EVSIISOMGRNRAKULLOSEPENB--LOETSLSVQSLNGLGVRLTRKORLOPKT 176
Db 526 ACGLGITLACSOERDISCPSSVDGTGSMPTTLDTATYK-FNAGLITLKNKR----- 577
OY 177 SVYIELGSDSEDTYKATYCSVGDDELQTPQGTROEISLDSAKKACERSETDV--- 233
Db 578 -----KFTY-SYSDASLQKKLQTHQLELTNL-SQOLENSAFEVLT 619
OY 234 -TTEHHQPSNNDLNTTEKRAERHPEKYQGSVSNLHVEPGCTHTHASSLOHENSU-- 290
Db 620 FTNVNSGIPDSSD-----KKRCLPNDE-----EPSTNSFGTATSKETSIYHA 663
OY 291 LTKDMANVEKAFCKSKQOPGLASQHNKMGSKETC-NDRTRTST--EKKVDLNA-- 344
Db 664 LISQD-LNDKEAIVIEEKQPYTAREADFLCLPRTGENDQSKPVSNGREKVLVSACL 722
OY 345 -----DPLCKERKMNKOKL-PCSENPRTGEDVPM----- 372
Db 723 PSAVOLSSISFSEQENPLCDHNGTSTIKLTPTPSKPLPSKADWVNSKEKKMKPEKLOESC 782
OY 373 ---ITLNSIOKYNEMFSRDELIGSDSDHSESESNKV-----ADVLVINEVDEY 422
Db 783 KVIIELSKNILVENEICILSE-----NSKTPGLPBGNIIEIVASSMSKQ 827
OY 423 SGSESEKIDLLASDPHALLCKSER--VHSKSVESNIEDK-----FGRTYRKASLP 472
Db 828 FQONAKI-----VIQDKQSPFISEVAVMNSSEELPPDSGNPFAQVTKCKNP 877
OY 473 NLSHVTENLIGAFVPEQIIQERPL--TNKLRKRRTSGLHPDEFKAKDLAVOKTPE 530
Db 878 DLGSSVE-----LOEEDLSHTQGSLSKNSPMANVEDVDAAHQAQVLTAKSD 924
OY 531 MINQGTQEQNGQVNNITNSGHEKNT-KGDSIQNEKNPNPIESLEKESAFKTYAEPIS 589
Db 925 SLAVVHDYTEKSRNNIEHQKQTEDEKFSNLSLMMKSGDNCSDSKWSEF--LDPVLN 981
OY 590 -----SISNMELELNHNKAPKK--NRLRRKSSR-----HIHMLELVSNLSP 634
Db 982 HNEGSEFRTASKKEIKLSEHNVKKSKMFKDIEQYPTRLACIDIVNLTPLANOKISEP 1041
OY 635 NCTELOIDSCSSSEELKKKYNQMPVHRHSRNQLMEGEPATGAKKSNKPNEQTSKRHDS 694
Db 1042 HIFDLK-----SVTTVYTSQHSSVSH-----EDT 1067
OY 695 DTFPELKITNAPGSEFTKCSNTSELKEFYNPSPILPREKEKELETYK-----VSNNADP 747
Db 1068 DTAPQW-LSSKODFHSNNLTTSQKAEITELSTILESGSQEFQFRKPSHIAONTSEVP 1126
OY 748 KQMLMSGEVLOT-ERSVSSSISLVPGDGTQGISLLEVSTLGRKKTPEPNKCVSQA 806
Db 1127 GNQWV---VLSASKWKMDTDLPLVPDSVQDTHSKQFQSGA-GVQSPFILLEPTCN 1181
OY 807 AEENPKGLIHGCSKNRNDTE--GFKYPLGHEVHNSRPTSIEMESSELDQYLQNTFKVS 864
Db 1182 --KNT-----SCFLNINEMERGFCALGTRKLSVSN-----A 1213
OY 865 KRQSFALSNPQNAEECA-----TFSAHSGS-----LKKQSPKVTFE-----CEQK 906

```

```

Db 1214 LRRKAMLFSDIENSEEPPSAKVGPRGFSASHHDSVAVFRIKIKKONTEKSEDEKSSKQVY 1273
OY EENO-----GKN-ESNFKPV-----TVNTAGEPVY---GOK 935
Db 1274 LÖNNIEMETICIFVGRNPEKTIKNTKHEDSTSSORNNLENSDGSMSSTSG-PYIHKGDG 1332
OY 936 DKPVD-NAKCSIKGSRFLCSQFRNGETGLTPNKHGLLÖNPYRIPPLPIKSFVYKTC 994
Db 1333 DLPAADGSKC-----PESC--TOYAREENTQIKENISDLCLFIMAEETCMKSSPKQL 1385
OY 995 KAMULEENPEEHSMSPREKMGNNINISTYSTIRNNIRENVFEASSNINEVGSSTINEV 1054
Db 1386 PSQKMONIKEFNIS-----FOTASGKNTR--VSKESLMSKNVININRREDE- 1429
OY 1055 GSSINEIGSSDENIQAELGRNRPKINAMRLGVLOPEVYKQSLPGSNCKHPEIK-KOEY 1113
Db 1430 ---LVIYISDS-----LNSKLHGINKDKMH-----TSCKKRAISTIKVF 1465
OY 1114 EEVYQTVNTDFSP-----YLISDNLEQPMGSSHAS-----QVCSETPD---DLDDGE- 1158
Db 1466 EDHFPVIVTSQLPAAQHPPEYEIESTKEPTLLSFHTASGKKVKIMQSLDKVKMLPDETQY 1525
OY 1159 IKEDTFAEND--IKSSSAVFSKSVOKGELSRSPSPPTHTLAQYRRGAKKLESSENL 1216
Db 1526 VRKTSFSGSGKPLKDSKKELTLAYEKIEVY-----ASKCEHQNFV 1567
OY 1217 SSDEDELPCQHLLFGKVNINIPQSTRHSVATE-----CLSKNN-----TEENL 1260
Db 1568 SKTEMLPOONTHMYKOTENLKTSGTSSKVOENIENNEKNPRICICQSSYPVEDSA 1627
OY 1261 LS--LKNLSNDSNOVILAK-----ASQENH- 1284
Db 1628 LAYTDSRRTCYRESLSGKRWLRQDGKLTGRTMTIKTECVKEHTEDEAGANASYEHL 1687
OY 1285 -----LSEETCASLFSQSCSEDL----- 1306
Db 1688 VIIRTEIDTNHVSQVSTLLSDPNVCHSYLSOSSFCHCDMHNDGFLKKNIDSDVP 1747
OY 1307 -----TAN-----TNQ-----DPLIGSK-- 1322
Db 1748 DMKNAGNTISPRVSAATKERNLHPOTIINEYCVOKLETNISPANKDYAIDPSLSDSNCK 1807
OY 1323 -----QMRHÖSÖGVLSDKELVSDDEERGTLLENNOEÖSMDNLGEASGESET 1376
Db 1808 VGSILVFITASHQETE---RTKEIVTDNCYKIV-----EQNRQSKPTQCTSCHKVL 1855
OY 1377 SVSEDCSGSSQSDILITQOROT-----MOHNLIKIQÖEMAELEAVLEÖGSGOPSNS 1428
Db 1856 DSKRDIICPSSSGDVICINSRKDSCPHNEQILQHN-----QSMS-----GLKKAAT 1901
OY 1429 YPILISDSALDELRLRPEO-----STSEKAVILTSOKSESEYPISONPEGISADK 1476
Db 1902 PRVGLTWTQTSKIRPPQAAHPSRTYGTJSTASGAIOVSDASLE-KARQVFSEMDGA 1960
OY 1477 FEVSADSTSKNEP--GVERSSPSKCPSLDRMYHSCSGSLQ-----NNYV---- 1523
Db 1961 KÖLSMWSLEGNKPRHVSXRENS-----VYHSTQGVLSLKPRLPGVNNSSVFCG 2010
OY 1524 -----SOEELIKYVDV-----EEOOLEEG--PHDLTETSYLPRODLESTPY 1563
Db 2011 FSTAGKRLVYSESALHKYGMLEFDLIRTEHTLQHSPIPEDVYSILPQCAEIRTPREY 2070
OY 1564 -LESGLISLPSDDESPSEDRAPASARVGNIPSTGALKVPOLKVAESQSPAAHTDTJ 1622
Db 2071 PVNSKILQKTYTNDKSSLPNSK--ESGSSGNTQSIIEVSLQÖMERODTOLVIGTVSHS 2128
OY 1623 AG-----YNAMEESVREKPEL-TASTERYNKRMSVWVSGTLPPEELVYKFAKHHTIL 1676
Db 2129 KAMILGKEQTLPNIKVKTDEMTFSDVPYKTVNGEYSEKESNEYFETAESAKAFMED 2188
OY 1677 TNLITETHTVHWKTADEPVCERTLKYFLGCIAGKVVVSYFWVYQSIKEKMLNEHD--F 1734
Db 2189 DELTDSÖTHAKCSL--FTCPÖNETLTFNSRTRKRGVTVDAVGOPPIKSLINEDRII 2245

```

```

OY 1735 EVRGDVVNGRNHOGPKRARES---ODRKIF---RGEICCYGPF 1772
Db 2246 ESKGSLIT-----PSKSTPDGYTKDRLSLFTHHMSLEPVTGCF 2283

RESULT 8
T14156
kinesin-related protein - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 11-May-2000
C:Accession: T14156
R:Wood, K.W.; Sakowicz, R.; Goldstein, L.S.; Cleveland, D.W.
Cell 91, 357-366, 1997
A:Title: CENP-E is a plus end-directed kinetochore motor required for metaphase chrom
A:Reference number: Z17893; MIMD:98028574; PMID:9363944
A:Accession: T14156
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-2954 <MOO>
A:Cross-references: EMBL:AF027728; NID:g2586070; PID:g2586071; PIDN:AAC60300.1
A:Genetics:
C:Gene: XCENP-E
C:Superfamily: centromere protein E; kinesin motor domain homology

Query Match 3.7%; Score 355; DB 2; Length 2954;
Best Local Similarity 19.3%; Pred. No. 1.3e-07;
Matches 358; Conservative 276; Mismatches 678; Indels 546; Gaps 87;

OY 65 KNDITKSLQESTRFSLQVELLKIICAFOLDTG-LEVANSYNFAKKE--NNSPEH---- 117
Db 552 EKELT--SLQQLQÖSKR--EKKELVQSFELKLALEEQSLVAKKNLEMTNSREHINA 607
OY 118 -----LKQEVSTIQSMGYRNRAKRLQSEPNPSIQETSLS----- 153
Db 608 EVQTDVEKEVYRKEMSVYLGDSGY-----NASMSDQÖDSSVDGKRLSSHDECIEH 657
OY 154 -----VOLSNLGVTRTLRTKQRIQPKTSYVTELGSSSEPTVKATAYCSGQDEL 205
Db 658 RKLMEQKIVLDEEF--TENLNKSSENDKÖKS-----SEDDFMSTIQLC-----EAI 701
OY 206 QITPGQTRDEISL--DSAKKACEFSET---DYTNTEHHOPSNNDLTTEKRAERPERK 260
Db 702 MARKANALFELALMRDNFDNILE-NETLKREIADRLERSLKENQETNEFEILERETOKE- 759
OY 261 YQSSSVSNLHVEFCGTHASSLQÖHENSLLTKDRNNVKAEPCKNSKQPG-----LA 314
Db 760 -----HEAQLIHETIGSL-----KKLVENAEVYNÖNLEEDLETKTKL 796
OY 315 RSQHNRMAGSKETCNDRTSTETKV---DL-----NADPLCEKKEWKKOLPCSEN-PR 365
Db 797 KEQEIQLAEIR-----KRAONLQKKVNFPLDSVMSGSEKICEITPOLKQSLSDAEAVTR 851
OY 366 DTED-----VPVITLSSIQKVNEMFSRSEDELGSDSHDGESESNK-----VADVL 413
Db 852 DAQKCEFSLESENLELKEKEDTSMNYÖKEKAASLFEKQLETERKSYKKMEADLQÖELQ 911
OY 414 DVNLNEDVYSGSSEKIDLLASDPHEALICKSE-----RVHSKYVESNIEDKIFGTYRK 468
Db 912 SAFNEINLYNG-----LLAGKVPRLDSRVELEKKVSEFSKQDEKALEEK-----NALENE 962
OY 469 AS-----LPN-----LSHYTENLILGATVTEPQIIQÖEPRTLTKLRKRRTSGHL 513
Db 963 VTLSEYKFLPNVEYECIKNOISKASEITML-----LQÖE-----GEH 999
OY 514 PEDFIKADLAVQKTEMINQÖTNO--TEONGOVYMNITNSCHENKTKQDSIQNEKNPPIE 572
Db 1000 SASITISQÖETIMQÖSQIQLQTLDEVYTHQSQVOQFEQYLEMKKHNDL----- 1049
OY 573 SLEKESAFKTAEPRISSISNME-----LELNTHNSKAPKKNRLRRKSSRTHIALELV 627
Db 1050 -FEKYIRNKSSEADLLREMNELKGTMESEVYKLAOTRKHELEETIRKQÖULLH----- 1100

```

```

QY 628 SNNLSPNCTELQIDSCSSSEETKRRYNQMPVYRSHNLQIMEGKEPATGAKKSNKPNQ 697
Db 1101 -----EKKY-----FEOAQOTIPPIPL----- 1118
QY 668 TSKRHSOTFFPLKLTNAGSFTKCSNTSELKEFVNPSPLPREKKEKLETVKYSNNAEDP 747
Db 1119 -----SDSLPSPKIVEG-----NSODPIEINDYHMLIALATERNNIMCLETERRSLKE 1167
QY 748 KOLMLSGEVLQTERSVSSSISLVPGTGYQTESILLEVSTLGAKEPBNKVCSCAA 807
Db 1168 QVIDANTQLOAQISQIEKSDIQ-KKQDLEBEVALLLEME----- 1208
QY 808 FENPKGLHGCSKDNKNDTEGFKYPLGH-EVNHRSRET-SIEMESELDAQYLQNTFRVSK 865
Db 1209 -----LKGHILTDSOLSTIE-KLOLENLEVEKTLQTOEEMKNITIERNELQNTFEDLK 1260
QY 866 ROSFALFSPGNAEECAFFSAHSGSLKQSPKVTETECQKEENQCKNESNIPVOTVIN 925
Db 1261 -----AEHDSLKOD-----LSENIQSITETODE-----L 1284
QY 926 TAGFPVVGQKDPVDNAK-----CSIKGGSRCFLSQFNGNETGLTPNKHGLQMPYR 980
Db 1285 RAAQELREKQKOLVVSFRQOLDCSV-----GIESPN-HDAVANOEKY 1326
QY 981 PPLPFIKSF-----YTKCKCKNLE-ENFEHSMSPREKNGENIPSTVSTISR 1028
Db 1327 -SLGEVNSTQSEMLRGERDELQTSCKALVSELELLRAHVKSVEGE--NLEITKRLNGLEK 1383
QY 1029 NNIRENVFEKASSNINEVGSSTNEVGSISINEGSSDENIQOELGNRQPKLNAMLRLGY 1088
Db 1384 ELTGSESESEVYKSMLENKEDNNKLEKQAEVSKENOPSLFEVSGSKL-----Y 1436
QY 1089 LQPEVYKOSLPSSNCKHPEIKKOEVEVQVNTDFSPYLLISDNLQPMGSSHASOVCE 1148
Db 1437 DEIEVLKAKOLKAAE-ERLEIKRDYFELYOTANTN---LVBGKLETPLOADHE---ED 1487
QY 1149 TPDDLLDGEIK-----EDTSPAENDIKESSAVPSKQKGELSPPRFTTHLAQGY 1202
Db 1488 SIDRSEMEIKVLGEKLEKERNQYLLERLOEKELELSNKLEILOKEMETSVLLKDDLOQ-- 1545
QY 1203 RRGAKLES-----SEENLSSEDEDELPCFOHLLFGKVNINPSSQ 1241
Db 1546 -----KLESLSLENILKENIDTTLKHSHTQAOLOQTOELOAKNLIAASDNCPI- 1598
QY 1242 TRHSTVATCELSKNTENMLSKNSLNDSCN--QVILAKASQEHNLSEETKCSA-----S 1294
Db 1599 TOEKETYSADCVAR-LEEKILLTLELHQKTNEQKELHKEKNELEQAOVLEKCEVEHILMS 1657
QY 1295 LFSSQCSLEDLTANTNODPPLIGSSKOMR-----HOSQGY 1333
Db 1658 MIESK-SLESLOHEKHDTEDQLLAKQOQVVOYTOEKKELQOTHEHTLAEDVHLKENIEL 1716
QY 1334 GLSDELVSDDERETG-----LEENNOEPOS-----MDSNLGEASGCESTSVS 1379
Db 1717 GLNFR-----NEAQKTKTEQCLLENKLEQESQHRLOCEIEBLMSLKDKESALETIKESB 1773
QY 1380 EDCSGLSSQSDILITQ-----QDRTMQHNL-----IKIQOEAELFAVLE 1419
Db 1774 QKVINLQEMEMVLEMELEKNSQRTVIAERDQLODLRESVEMSIETODDLRKAQEALO 1833
QY 1420 OHGS--QPSNVPSTIISDSAL-----EDLRNPEST--SEKAVLT- 1456
Db 1834 QOKDKVQELTSQISVLOEKISLENNQMYNATVKEITLERSDLNOSKQHLESETTLTSL 1893
QY 1457 SOKSESEYPISONPEGLSADKFEVSADSTSKNKEPGEVRSPPSKCPSLDDRY-----MS 1512
Db 1894 SLKEKEFALQO-----AEKDKADAARKTIDITEKISNIEBOLLOQATNLKETLYERSLQ 1949
QY 1513 CSGSLQNNRNPQOELLIRVYDVE-----EQOLESGPH--DLTE--TSYLRQDLEGPYLE 1565
Db 1950 CKEQDLA-LNTEHLRRTLAKSKDALGKMQEOREANANKYALATEKMSLSLEQINENTVTLK 2008
QY 1566 SG-----ISLFSDDPESDPESEDRAPESARVGNIPSSTSALKVPQLKVAESAQSPAHAHTTD 1621

```

```

Db 2009 ECEGEKETFYLPORPSKQOSSQOME-----LRESIKTKDLOL-EEAKEISEATNE 2058
QY 1622 ----TAGYNAMESSVREKREBELT-ASTERVKNRMS--MYVSGLTPEEPMLYKKFARKH 1672
Db 2059 IKMLTAKISLEEEIIONASILNEAVSERENLRHSQOOLVSEL--EQLSTFLK-SRDH 2113

RESULT 9
F90073
hypothetical protein SA2447 [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: F90073
R:Kuroda, M.; Ohba, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; O
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K
C:Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiratazu, K.
Lancet 357, 1225-1240, 2001
A:title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: F90073
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-2271 <KUR>
A:Cross-references: GB:BA000018; PID:g13702612; PIDN:BAE43752.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: SA2447

Query Match 3.5%; Score 342; DB 2; Length 2271;
Best Local Similarity 18.5%; Pred. No. 3.3e-07;
Matches 310; Conservative 266; Mismatches 835; Indels 266; Gaps 49;

QY 161 TTTTLTKRIPORPOTSVYIELGSDSEEDTVKATYCSGDOELQITPQGTREISIDS 220
Db 638 TTTVSTVDANKKSTTTTIN-VDTAPTIV-----PIGQSSSEVSP-----ISP 683
QY 221 AKKACESETDYVTEHHOPS-----NNDLNTTERRAERHPEKYOGSSVSNLHYEP 273
Db 684 IKIATQDNGNAVTVMTVGLPGLTFDSNTNTISGT-----PTNIGTSTISIVTDA 735
QY 274 CGTNTHASSLOHENSLLTKRNVAKAEFCNKKQOGLARSOHNRKAGKEFCND--- 330
Db 736 SGNKT-TYTFKYE-----VTRNSMDSVSTSGSTQOSQSVSTKADQSASTSGSIV 789
QY 331 RRPSTTEKVDLN-ADPLCEKEMKKOKLPCSENPRDREDVPMITLNSIO-KVNEFPR 388
Db 790 STSASTSKSTSVLSDSVASAKSLSTSESNSVSSSTSVLNSQSVSSMSGVSKSTSL 849
QY 389 SDELLGSDSDHGESESNKAVADVLDVLENDV-----YSGSEKIDLLASDPHEALICK 443
Db 850 SDISINSNSTERKESLSTSTSDSLRTSTLSLSDSLMSSTGSLSKQSLSTIS- 902
QY 444 SERVHSKSVESNIEDKIRCKTKRKASLPNLSHVTEMLITGAFYEPQIIOERPLTKL 503
Db 903 ----GSSSTVSALST--SNAISTSTSLSESASTSDSISTNSIANSQSAS--TSKSD 953
QY 504 RRRPPTGLHPDEDFIKADLAQVOKTPENINQNTQONQOVNNTNGNGHKNKTKQSDIO 563
Db 954 SOSSTIS-----LSTSDSKMSMTSESLSDSTSGSVSSGLSIASQSVSTSDSMS 1006
QY 564 NEKNPNPIESLEKSAFKTYAEPISSISNNWELELNHNSKAPKKNLRKSSTRHIIHA 622
Db 1007 TSEIVSDSISTSGSLASDSKMSVSSMST-----SQSGSTSESLSDSQSTSDS 1058
QY 623 LELVYSRNISPNCELOI-----DSCSSSEETKRRYNQMPVYRSHNLQIMEGKEPA 675
Db 1059 KSLSLSTSGSTSTSTSTSTASVRTSESQSTSGSASQSDSMKISTSTSTSDSKSAS 1118
QY 676 TGAKKSNKPNQOTSRRHSDPPELKLTNAGSPFKCSNTSELKEFVNPSPLPREKKEKL 735
Db 1119 TASSBSISQASASTSGSVSTSTSLSTNSERTSTSVSDSTSLSTSESDSI--SESTST 1176

```

```

QY 736 ETWVKNNAEDPKDMLSGERVLOTERVSESS---SISLVPCTGTQTESISL----- 786
Db 1177 DSIISAISASESTISLSESNSTSDSESSAGAFISELSESTSESSVSSSESS 1236
QY 787 ---EVSITLKATERNKCVSOCAAFENRGLIHGSKDNKRNTEGFKYLGEVNHRSRT 843
Db 1237 LSDSTSESSSTISLSTNSGASISLSTISSESTN-PEKSESVSTLSMSTISLST 1295
QY 844 SIEMEESLDAQYLONTFKVSKROSFALFSPNGNNEECATSAHSGSLKQSPVTFPC 903
Db 1296 SILSTLSBDSKSDSLSTMSLSTISLST---SKSDSLSTISLSTSGSTSES-----ES 1346
QY 904 ECKEENOGKESNIRPVQTVNITAGFPVYGKDKPDVNAKCSI-----KGSFECULSS 956
Db 1347 DSTSSSEKSDSTMSISMSOSTSTSTSTSLSDSTSLSLASMNQGVSNMS 1406
QY 957 QPRGNETGLITNKGILLONPRIPPLPKSFV-----KTKCKNLEENFEHSHS 1009
Db 1407 QASASTSTSTSESSQSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 1466
QY 1010 ---PEREMGNENIPSTVSTISNNIRENYFKASSNINVEGSSST----- 1051
Db 1467 LSGSESSQSTISASSESTSEASTSLSDSTSTNSGASTSTSLNSASASESDSS 1526
QY 1052 ---NEVGSSINEIGSSDENIOAELGRNGPKL-NAMLRGLVLOPEVYKOSLPGSNCKHP 1106
Db 1527 TSLSDSTASQSSQSTSTASISLSDSLSTSTSNRMTIASLSTSV-STSGSTSEST 1585
QY 1107 ELKKEFEVVOVTMTDPSPL-----ISDNLEDPWGSNHAQVCS 1147
Db 1586 SESDSTSTSLSDQSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 1645
QY 1148 EPDDLDDGELEKEPTFAENDEKSSAVFSSKVOKGELSPSPFTTHL----- 1198
Db 1646 STSTMSDSTSLSDVSDSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 1702
QY 1199 -AGYRGAKKLESSEENLSED-----BELPCFOHLECK 1233
Db 1703 DSQMSSESVNDESSSESSSESSKMSGSTSVSDGSLSVSTSLRKSSESVSSSLSS 1762
QY 1234 VNINISQSTRHSTVATECLSKTEENLKLKNSLNDCSNOVILAAQOEHLSEETKCA 1293
Db 1763 QMSDVSSTSDSSLSVSTSLRSES-VSESDLSBKSTSGSTSTSTSTSTSTSTSTSG 1821
QY 1294 SLFSSQCELEDD---LTANTNTQDP-----FLIGSSKOMRHOSEQVGLSDKEIVD 1343
Db 1822 SESVEESTSLSDISMSDSTSTSDSLSGSTSLSTSDLSBKSLSSQSSMSG 1881
QY 1344 DEERGTLGEENNOEBS---MDSNLGEAAGCESESTSEVSDSGLSQSDILTTQORDTM 1400
Db 1882 SESSTSVSDQSSSTSNQFPMISASSEDSMGTSDOSNLSGNSSTSTSTSTST--DSM 1939
QY 1401 QNHLIKLOQEMAELEAVLEOHGOPSNTPTIISDSSALEDLRNPDESTSEKAVLTSOKS 1460
Db 1940 SGSV-----SVSTSTSLSDISGISTSVSDSSSTSTSTSLSDMSQSGSTSTASGSLST 1994
QY 1461 SEYPTSONPEGLADKFEVSADSTS-----KKEGCVRRSSPKCPSLDDMY 1509
Db 1995 ISTSMASSTSSQSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 2054
QY 1510 MH-SCSGSLQNRNYPQOELLKVVDVEQOLESQPHDLTERTSYLPRODLESTPYLESQ 1568
Db 2055 LSTSDSDSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 2107
QY 1569 SLFSDPDSESDAPESARVGNTPSSTSAKLVQKVAESAQPAANHITDTGYNAM 1628
Db 2108 SMSQKSDSTISDE-----SVSTSTST---SLSTSDST-STSESLSTSMGSGSOSI 2155
QY 1629 EESVREKRELTASTERYNKRMSVYVGLTPEEFMUYKFAKKHITLITLITEET 1685
Db 2156 SDSTSTSMGSGSTSES-----NSMHPSDSMH---HTHSTSTSLSSSEATT 2200

```

```

RESULT 10
T41023
probable nuclear pore complex-associated protein - fission yeast (Schizosaccharomyces
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence-revision 03-Dec-1999 #text-change 03-Dec-1999
C:Accession: T41023
R:Murphy, L.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, June 1998
A:Accession number: Z21965
A:Accession: T41023
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1837 <MUR>
A:Cross-references: EMBL:AL023860; PIDN:CAA19588.1; GSPDB:GN00068; SPDB:SPCC162.08C
C:Genetics:
A:Experimental source: strain 97zh-; cosmid c162
A:Gene: SPDB:SPCC162.08C
A:Map position: 3

```

```

Query Match          3 5%; Score 341.5; DB 2; Length 1837;
Best local similarity 19.4%; Pred. No. 2.6e-07;
Matches 351; Conservative 310; Mismatches 674; Indels 475; Gaps 78;

```

```

QY 2 DLASLVEEVQNVINAMOKILECPI-CLELIREPVSTKCDHIFCFMKILNOKKPSQ 60
Db 130 DQNLKQTAQLNLLSDKEVEKEKITLIKDALASSHQV-----LELQHTQO---- 179
QY 61 CPLKNDITKRSLQESTRPSQVLEBLKTKCAFOLDTGLEVANSYNFAKE--NNSPEHL 118
Db 180 -----EKASLQNYEP-----ELQKLT---QKNSILNNTWLSRELQGVNDKLSTL 223
QY 119 KDEVSIQSMGYRNRKAKRLQSEPPNSLOETSLVSQSLNLG---TVPTLR-----T 167
Db 224 HOEASLEKSO-----LSSQLDAVLEKALQKKSLSQQTESLRLPONIYALS 274
QY 168 KQRIQPKTSVYIELGSDS---SEDTVKATYCSV-----GDOELQITPOGTRD 214
Db 275 EMRKQYEFQVFEKEISSQKQISELMMEKCDCSLRLELQNSGELKLELAQSSPE 334
QY 215 ELISDSAK-----KAACEFSETDVNTEHHPNSNNDLNTTKRAERPEKYQSSSVN 268
Db 335 E-QLESHEKAEASLSQINFLKEVSSLE-----SOLKANERL--RHYDEIETSDME 385
QY 269 LHVEPCGNT-----HASSLOHENSLLTTRD--MNVEKAFECNKS-----KOPG 312
Db 386 LKYSMLNLSMKGFGQSSVDLYSERLYKKQKQTCQEVRLORSTYHWEANLOHP 445
QY 313 LARSQHNWAGSKE---TCNDRRTPESTKRYDLNADPLCEKEMKMKQL-----PCSEN 363
Db 446 LKEDQFKRAHQREIVAMSEQYKSLD-----QKAKSYEQLETLFKKKTEN 496
QY 364 ---PROTEDVP-----WITLN--SSIQKVNEMFSRSDDELGSDDSGE 404
Db 497 KHYEOTKDLARQOVOLLHELDELICENGIVLVGDSKRTKINSYEKS---LTBEDPTDOI 553
QY 405 SNAKYA-----DIVDLVNEVDEYSSSSKIDILASDPREALICKS----- 444
Db 554 SRLVYFRNIRLQOONOLSAVHELADRMKEDKPDLDGAEIOEETLIKANETIDOLT 613
QY 445 ---ERVHKSYSVSNTEKIFGKTYRKKASLPMLSVHTENLIGAVTEP-----OIIOER 496
Db 614 KMLEVSDQLKRSLSKERPFRLVQENKELDMAPATPNSKNTNLIDETSYQRLIRE 673
QY 497 PLTNLKKRRKPTSLGHEPDKKADLAVOKTPREMINOCTNOTEDONGOVANITNSGHENK 556
Db 674 QLTNLELSKISIRNKE-----KPEEAISSL-----OLEKSNTOLOLTSLTSR- 718
QY 557 TKGDSIQNEKNPNFEST-----EKESAKTKAEPDIS--SISNMELEANTHNSKA 605
Db 719 ---SLAEKLNDELKSLVLSRSDDELDESKLSQEOBASKKTFYOVVSSOLSLCNSOL 774
QY 606 PKKNRL--RRKSTRHIALELVSRNLS--PPNCTELOIDSCSSSEIKKKKYNQMBVR 661

```

```

Db 775 EQSHIYDNLKSENLLTTSVKOKLKAADLSNLSKSLSSLOQDNFHKAKQIETESSNOEYATV 834
QY 662 HSRNLOIMEGKEPATGAKKSNKPNEQTSKRHD-SDTPEELKLTNAGSF-----T 710
Db 835 DSMNRILE-----LSNDLRVANKSLSCSDVARTLQN---SDTLREHOTLVQ 882
QY 711 KCSNTSELKEVNPDLPR-----EKEEKLTVKVSNNADPKDL 750
Db 883 LQSNITTELKODI--TLQRTVRNOLEIQTTELKERLKFMEERENQSKITIAN---KDT 936
QY 751 MLSEERVQTERSVSSISLVPGDYGTOSSISLSEVSTGKATEPDKVCQSCAAEN 810
Db 937 TQNDNV-----EVAISIEL-----ERTKEKLMAE--LEKSIQOXYLASE----- 977
QY 811 PKGLIHGSKDNNDT-EGFYRPLGHEVNSRE-TSIE-----MESEIDA 854
Db 978 -KTL-----EMMNETHQFHLVSEISTREKITSLSRELLDKRKREVLKEKESSS 1030
QY 855 QYLONTKRVSKRQSFALFSPNGNAEECATSAHSGLKQSPK--VFCECEQKEENO- 910
Db 1031 KELAKOLEDAVREKDSALSFKKDKYKIRSDADRVYTSLEKEDIKERSLMKECHSNYESEI 1090
QY 911 ---GKNESNIKPYQT-----VNITAGFPVY-----GQKD-----KPYDNKAKSIR 947
Db 1091 VSHGRTTQKLDLRTEDDEVNTKYTLKLANFEQOHSGLSGAEKDNINORKAMEDITSILK 1150
QY 948 G-----GSRFCLSSOF-----RGNETGLITPNKGLLQNPYRIPPLPIKSF 989
Db 1151 DYILGLENONKLLHSDPDLSSQOITVLOONSENENITANLEAVQDNDR-----ELVSY 1205
QY 990 VKTKCKANLEENEEHSMSPREKMGENIPSTVSTIRNNIRENVEKEASSN----- 1043
Db 1206 LRH--EKEIMDNKY-ELTILDRGL-NOOVKSLQSTVSLQLELRLQSLPVSNOQTDP 1261
QY 1044 -----INEVGSSTNVGSSINEIGSSDENIOAEIRN 1075
Db 1262 IISGSOEVLIESNSVLRKNDQAKLGITQLEKEVEKELNLSLNPLOETINELKKEITAK 1321
QY 1076 RG-----PKLNMALRLGLVLOPEVYKQSLPGSNCKHBEIKQOEVEYVQ 1118
Db 1322 TASLNLMEKYNRMKLRFOVLYNKERYVDPQLEELK-----NCEALEKEKQOLETQLQ 1376
QY 1119 TVNIDFSPYLSLDLEQPMGSHASQVSETPDOLLGDEIKEDTSPFENDIKESSAVFS 1178
Db 1377 ETAKE-----TDFKQOVNS-----LNEEVENL-----KKEVEQANTKTRLAAMVN 1418
QY 1179 ---KSVQGEISRSPFPTHHLAOGYRGAKKLES---SEPNLSSEDEELPCFOHLLFG 1232
Db 1419 EKCENLKKSILTR-----FAHLKOELTNKKKELTSKAENEMOKETIESLKDSDHQLQ 1472
QY 1233 KVNINPQOSTRHSVTATECLSKNTEENLLSKNSLNDGCSNOVITAKASQE-----HH 1284
Db 1473 SASDADQOITKEQFEQLEKSEKEREKELADSKNELEHQSEAVDADGKTETISNLEKETHE 1532
QY 1285 LSEETKCSASLFSQCSLELTLTANTNTQDPF-----LIGSSKQMRHOSSESQV 1333
Db 1533 LRSQKEGLVOOVQMLSAELALRHSPTQSGLENADLARLSQLESTKQYEEKERETEI 1592
QY 1334 GLSDKEGLVSDDEEGTGLEENNOEQSDSNLGEAASCESETSVSEDCSGLSQSDSLT 1393
Db 1593 LAARSELVAEKEKKELEENLNEKSOIKELQEAQKNSSEN-----HDNDDDTIK 1645
QY 1394 TOORDTMOHNL-----IKLOEMAELE-----AVLEQHGSOPSNSPYIISDSSAL-ED 1441
Db 1646 QOYVEEKLEKENSANDVLLKVVVAETEPFSKAKISYEKKTRDLONKITQLEETLENLKO 1705
QY 1442 LRNPQOS-----TSEKAVLTSQKSSFTPISONPEGLSADKFEVYADSTSK-----NKE 1490
Db 1706 LSNPEKDESTSVTETKPYTSKPTASKADVQGNATEASSAKRBPSCGSLARLQGTGKO 1765
QY 1491 PGEVRSPPK 1500
Db 1766 KGVORPAVSR 1775

```

```

RESULT 11
JC5837
364k Golgi complex-associated protein - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 05-Mar-1998 #sequence_revision 13-Mar-1998 #text_change 20-Jun-2000
C:Accession: JC5837
C:Tokl, C.; Fujiwara, T.; Sohda, M.; Hong, H.S.; Misumi, Y.; Ikehara, Y.
Cell Struct. Funct. 22, 565-577, 1997
A:Title: Identification and characterization of rat 364-kDa Golgi-associated protein
A:Reference number: JC5837; MUID:96093490; PMID:9431462
A:Accession: JC5837
A:Residues: 1-3187 <PRO>
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Cross-references: DDBJ:D25543; NID:9516825; PID:BA05026.1; PID:9516826
C:Comment: This protein plays a role in the formation and maintenance of the characte
C:Superfamily: giantin
F:49-549,624-1176,1238-1707,1763-3114/Domain: coiled-coil leucine zipper #status pred
F:3165-3187/Domain: membrane anchor #status predicted <MAD>

```

```

Query Match 3 5% Score 340.5; DB: 2; Length 3187;
Best Local Similarity 18.9%; Pred. No. 6.2e-07;
Matches 375; Conservative 326; Mismatches 719; Indels 561; Gaps 91;

```

```

QY 23 ECPLELTKPVPSTKCDHIFCKFCMLKLLNOKKGPSCQPLCKNDIKRSLOSTRPSOL 82
Db 919 EEPVCKEALQOLE-----WLRKESQRRKRIQAL--ISKRELLQ--KYSKL 962
QY 83 VEELKITICAFOLDTGLEVANSYNPAKKNNSPHLKDVEYSIIQSMGRNRRARLLQSEP 142
Db 963 EEELAKV-----RESETKDSL-----RESEKRELEDS 990
QY 143 ENP-----SLQETSLSVQSLNLTGVTFRTKQRIQPKSVYIELGSDS 186
Db 991 KKKDDPEKVTGSEWRELEVSRLTISEKEVELEIRDLKKAABEELQALVQRM--- 1046
QY 187 SEDTVNKAATYCSVQDQLL--QITPQ-----GTRDEISLDSAKKACFSETDVTNTHHQ 240
Db 1047 TQDLQNKTKQDLLELTETENQATIQKFTGTMTAGGDSAVKETSVPFRAGGEGHK 1106
QY 241 PSNNDLNTTKRAARHDEKTYGSSVSNLHVEPCGTHTASSLOHENSLLTKDRMNE 300
Db 1107 PE-----LECKIVDLKEKTYQ-----LOKKQLEALLISRKAI-LK 1139
QY 301 KAERCNK-----SKQGLARSQHNRAAGSKETCNDRRPTSTKKVLDNADPLCEKEMN 354
Db 1140 KAQEKELKLEELKEQKQDADRHLQEDFGOSKENENIRAPLRO-----LOAKEST 1189
QY 355 KQKLPCSENPPTDEDVWITLNSSTQKVNEMFSRSDLELGSDDSHQES---ESNAKVAD 411
Db 1190 DQQLPGTQOQRPHTGSBGLSEGT-----EPASSD--LHMAQVSHQETITLQATYSVQ 1243
QY 412 VLDVLNVEDEYSGSSEKIDLASDPHEALLICKSERV----- 447
Db 1244 IQDOLKEIEV---EKEELEKISSSTSELTKKSEVLLQEQINQEOGLEIQLNKAASHBA 1300
QY 448 --HKSQVSNIEDIIPKQTYRKASLNL-----SYHTE-----NLITGAFVIE 489
Db 1301 KAHTEQKQLELE-----SSQIKTADLEHLTLQLELETQKHGQKEEVSYSVLVQGLER 1355
QY 490 PQIIQERPLTKLTKRRRPTSGLHP-----EDFIKKRADLV---OKTPEMINQGTN 537
Db 1356 EQTL--TTVQTEMEQERLIRKALHTQLEMAKKEHEBLKQVOYEICELKKQPELEESK 1413
QY 538 QTEQNGQVMNITNSGHNKTKGDSIQNEKNPNPTLEKE--SAFKTKAEPISISSIMEL 596
Db 1414 AKQO-----LQKIQALALISRKALKLEKNSIQLOELSSARDAVEHLTRSLADVS 1462
QY 597 ELNTHNSKAPKKNLKRKSSTRHHALELVYSKLS-----PPMCTELQIDSCSSSE 649
Db 1463 QVSVOQO---EKDALLKLLALQEERDKLIVEMDKSLLENOSLGSCCESLKLALGLTGD 1519

```

QY	650	KKKKKYNMPVRHSNNIOLAMGBEPAGCAKKSNNPNEDTSKRHSDTPEPLKLTNAPCSF	709
Db	1520	KEKLTKELESYRCSIAESTWQE-----KHKLOEY-----EVLLO-----SY	1559
QY	710	TKCSNTSELKEFVNPSLPREKE-----EKLEFVKVSNNAEDPKDMLSGERVLO	759
Db	1560	ENVSNEARIOHVVESVROEKOYVAKIARASBDRKEKOLQDAGCEMEEM--KEKKRK	1617
QY	760	TERVSESSISLVPECTDYGTQOESISL-LEVSTLGKAKTEPNKYSCQAFENPKGLIHQC	818
Db	1618	FAKSKOQKILFL-----EENENRLRAEAPVAGGANSMEALLSNASL---KEELERI	1667
QY	819	SKDNNDKEGKPYLGEVHNSRET-----SIMESESLDAQLOLNFK-----	862
Db	1668	TLEKTLSEKEALMAEKNTISEETRNLIKLOEAEOELQASLETTEKSDPKVILEEYTE	1727
QY	863	--VSKROSFALSPNGNAEEBCATFSAHSGSLKROSPKYTECEOEKENOGKNSNIXPV	920
Db	1728	AVWGSKOBDLSJENAKLEDAEATLLANSA---KPGVETSESHDINNYLOLDQIK--	1782
QY	921	QTVNITAFPVVQOKDKVDNAKKS IKGSGFCLSSQCRGNETGLITPNKGILLONPRTI	980
Db	1783	--GRIALEKEKKDRFL-----SOQLNE-----KALLTQ-----	1812
QY	981	PPLPEPIKSFVTK--CKKNLLEENEEHMSPEREGNENIRISTVTSIRNNIREVFEKA	1039
Db	1813	-----ISAKOSELKLLLEEPAKIMML-----NQIOBELSRVTK--LKETAEEK	1855
QY	1040	SSSN---INEVGSSTNEVGSSTINEGSS--DENIOAELGNRGPKLMLRGLVQEV	1093
Db	1856	DDLEBERLMOJLAELNGSIGNYOQVDTVAQIINNEOLESEMON-----LKRCVSELEE	1906
QY	1094	YKOSLPGSNCK--HPEIKKOEEYVVOVNTPEFYSLLSDNLEOPMGSSHAOVCEPDD	1152
Db	1907	EKQOLVKEKTYVESEIRK-ETMEKIQAGQ-----KGGESKIHAEEL-----QE	1948
QY	1153	LLDD--GEIKEDTSFAENDIKESSAVFSYOKGELSHSPSPFTHTHLAOYRGAKKLE	1210
Db	1949	LLKEKQGVKLOKDJCIYRLGRISAL--EKTVKALE-----EBOKLD	1993
QY	1211	SSEENLSEDE-----ELPCOHLLEGKNNIPSOSTRISYATEC--LSKNTEENILS	1262
Db	1994	ATKGNLAOVZHHKKAQOELSEFKLL---DJOSEKAR--VLADNIKIKELQSNES	2047
QY	1263	LKNSLINDCSNOVI--LAKASOEHNLSSETKCSASL--FSSQSELEDITLANTTODPFLIG	1319
Db	2048	IKSQIKQKDEDLRLRLEQAEKHKKKNMGEXKIDALHREKHAHBDTLAEIQVS---LTR	2104
QY	1320	SSKOMRHOSEOGVGLSD-----KELVSDEERGTCLENNQOEDOSMDSNIGEAAGCES	1374
Db	2105	KDKOMKELQOGLDSTLQALAAFTKSMSLDODRVIDEAKKWO---PRGDAIOTKEE	2160
QY	1375	ETVSJED--CSGLSSQ-----SDILITQOORDIMQHN---LIKLOO	1409
Db	2161	EVRLKEENCTALKOQLKQMTIHMELKITVSRLHDKHEIWSKATOLQHOQKAYDKLOE	2220
QY	1410	EMAELEAVLOHGS---OPSNSYSIISDSALE---DLRNP-----EOSTSEKAVLTS	1457
Db	2221	ENKELMSLOEAGLOLHDSKKNELTKLESELKSLDQSGDITLKNSELEKCREHNNEGITIK	2280
QY	1458	OKSESEYPISONDEGLSADKFEVSADSSRSKKKEKGVESSPSKCPSLDDRRYHMSCG--	1515
Db	2281	QEAIDQMKFCDELEND--LTASSELITRLHDELINVE--OKIITSL-----LSGKE	2328
QY	1516	-----SIQNRNYPQSOELLKIVVDYEEQO---LEESGPHDLTETSYLPRODEGTPYL	1564
Db	2329	EAIQVAINAEHQOSKEIKELLENLLOEEBENITLLEENKRAVEKTNQL-----TEAL	2381
QY	1565	ESGJSLSPDDPESDPSEDRAPESARVGNINIBESTALKAVPOLKVAESQSPAAHTTTTAC	1624
Db	2382	ET-----TKSESLEOKAOLDSFYKSMSSLODDDRRI-----VSD	2415

```

OY      1625  YNAME-----SVSRKPELTASTERYNKRMSVSGLTFEEMFLVYKFARKHITLTNLIT 1681
           ||| |||| | : : : ||
Db      2416  YRQLEERHLSTILEKDELIDQAAENNNLKKEIKRL-----RGH---MDDLNS 2466G
                                           ||| |||| | : : : ||
OY      1662  EETHHVAKTDAEFV-CERTLYKFLGIAGRWVSYFWYTOSIKERKMLN--EHDFEVR 1737F
           ||| |||| | : : : ||
Db      2461  ENA-----KLDAELIQYRRDINEVITI-----KDSOOROLLFAQLQONKREL 2502Z
                                           ||| |||| | : : : ||
OY      1738  GDVVNGRNHRQPKRKARESDQRKIRGLEICCYGPTTNRPQTGLEMMVLCCASVYKELS 1797F
           ||| |||| | : : : ||
Db      2503  NECV---KLEGRLLKGSEAEKOSL-----QMSLDALOEENO---GLSKETKS 2542Z
                                           ||| |||| | : : : ||
OY      1798  F 1798
           |
Db      2543  F. 2543

RESULT 12
A42771
reticulocyte-binding protein 1 - Plasmodium vivax
C.Species: Plasmodium vivax
C.Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 28-Apr-1995
C.Accession: A42771
R.Galinski, M.R.; Medina, C.C.; Ingravallo, P.; Barnwell, J.W.
Cell 69, 1213-1226, 1992
A.Title: A reticulocyte-binding protein complex of Plasmodium vivax merozoites.
A.Reference number: A42771; MUID:92315338; PMID:1617731
A.Accession: A42771
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1*2829 <GAL>
A.Experimental source: Belem strain, merozoites
A.Note: sequence extracted from NCBI backbone (NCBIN:108114, NCBIF:108115)

```

[illegible]



```

QY 425 SSEKIDLL-----ASDPEALICKSERVSHSKSVESNIEDKIFGKTYRRKASLPNLSHYT 478
D 1065 IKRDVDELNVNQVISEKNVTLFFKNNSVYIEAMHSHTVVAHGITSNKNELIKSKEVE 1124
QY 479 ENLIIGAVTEQIIQIÖERPLTKLRKRPRPTSGLHPEDFIKADLAVOCTPEMINOCTMO 538
D 1125 DKL-----NIVEONEDYKKRVKNPEN-----EKÖLEAIRGSMSTLKEVINHVSE 1168
QY 539 TEONQVANNITNSG-----HE-----NKTGCD-----SIOEKNPNPESLE--K 576
D 1169 MTOLESTANTLKSNAKKENENDLEBLNKTQOMRDITYEKLAKIAELKEGVNELKDAN 1228
QY 577 ESAFKTAPPISSISNNELELNHNSKAPK-----KXRLRR 613
D 1229 EKANVEPEPERNIIGHYLERITVEKDRAGKVEEKNSLKTKIEKLQETSDOSQELVYT 1288
QY 614 KSTPFIH-----ALELVSNLSPNCTELQIDSCSSEIK-----KKKYNQMPVH 662
D 1289 TSTKHLNNAKQYEDVIRKN-----EEDSTOLREKAKSLETLDKMKKLVQVY---- 1335
QY 663 SRNLOAMEGEPATGAKKSNKPNE-----QTSKHDSDTFPEELKTNA 705
D 1336 ---NNMLOAIOGNATISK---ELNELKGYIELLISTVSYSSILEYKKNSESVRFSQLAN- 1390
QY 706 PGSETKCSNTSELKEFFVNPDLPREKEEKLFTVKNVNAEDPKDMLSGEVLQTERSYE 765
D 1391 -GEFFKAGC-----BEKNASARLAEAEKLEQIVKDLDYSDIDDKYKKEGIRKELIKK 1444
QY 766 SSSISLVGCTDQCTÖESISLLEVSTLGRKAKTEPNKCVSOGAFENKGLIHGSKNRND 825
D 1445 ESALTVEESEKFKOMCSSHMEKNAKGGKRIEYL-----NN 1481
QY 826 TEGFYPLGHEVHNSRETSIEMEE-----SELDACIYQNTFYKSKROSFALFSPGNABE 880
D 1482 GDGGR-----ANITDSQMEVGNYSKAEHAFHYEAOYDKTKAC-----E 1523
QY 881 ECATFSAHSGSLKKOS--PKVTFECEQKEENQKNESNIKPVQVNTAGFPVYGQDKP 938
D 1524 SIVAVYTKMDNLFNESLMKEVYKCEKKDEAKYSAKLKP----- 1564
QY 939 VDNKCSIKGSRFLSSQFRNGENGLTPNKHGLLQNNRYRPLPFIKPKTKRKNL 998
D 1565 -----YDGRIRAVSENERKI-----SELKEKAKVE- 1590
QY 999 LEENEHEHSMBERMGENENIPSTVSTISRNIRRENVFEKASSNINEVGSSSTNEGSSIT 1058
D 1591 -----KRESSQINDVSTKSLQIDNCROOL--DSVLSNIGRKO-----NML 1630
QY 1059 NEIGSSDENIQ-----AELGRNRG--PKLNAMLRLGVLOPEVYKOSLPSGNCNKHPEIKOE 1112
D 1631 QYFDSADSKMSKSVLPISLGAESKSLDKVA-----AKES 1664
QY 1113 YEEVQVQVNTDPSFPLIDNLEOPMGS--SHASQVSEFPDILL-----DDG---EIKED 1162
D 1665 YEKNETLEVQNEKSR1-----NWE--GSLDIDDKITTDIENDLLKMKQYEEGLQDIKEN 1718
QY 1163 TSPFEND-----IKESSAVFSSVKO-----CELSRSP 1192
D 1719 ADKRKSNFELYGSELNALLDPTSTIFIKLKEKYDMTGKLYGYKAMNIEHFPKSYN- 1777
QY 1193 FTHHLL-----AQYRGAKKLESSEENLSEDEE-----LPCFQ 1227
D 1778 LIETHLSNATDYSVTFEKAQSLRELAEK---EEHRLRREBAIFLLNDIKKVESIKLK 1834
QY 1228 HLLGKVNINISQSTRHSTVATECISKNTENLISKNLNCOS---NQVILARASQ 1283
D 1835 EMMKRVSAEYEGMKRDHTSVSOLVODMKITVDELTKTLNDISECSSVLNNVSVIVKVKVS 1894
QY 1284 HLESTKCSALFSSQCELEDTANTNTQDPFLIGSSKOMRHQSPQ-----GV 1333
D 1895 KHADYRRAANSYEMSVT-----LANVPLSDEAKISSGMEFNAEMKSNFKTDLLEITSY 1949

```

```

QY 1334 GLSDKELVSDDEERGTELENNQEOQSDMSNIGEAASCESETSVSDCSGLSSQSDILTT 1393
D 1950 ISNSNELLKRTIÖQSDNDVYÖKERESEQLAKDQATDIYVYIKLNENEFNELEKAKNKEEYVS 2009
QY 1394 TQÖRQTMQ-----HNLIKQÖMALLEAVLDQHGQSPS----- 1426
D 2010 EKVREALKRLSQVEGIRCHFEFENHRLDTELEMLKMKVTTIYRDKSERESSQLOEMENE 2069
QY 1427 -NSYPS-----IISDSSALEDLRNPEOST-----SEKAVLTSOKSSE-----Y 1463
D 2070 MMTYSNITQLEGIVYVSAQESKEDIEKLENSNEERKNISEKISTSDSVIEMNSTIDELY 2129
QY 1464 PISQNDGELADKFEVVSADSTSK-----NKEPVERSPSKC-----PSLDRMYVHS 1512
D 2130 KLGKNCQAHMISLISYANMKTSKRLIMNK-----EKENTEKCVADYIKDNSSSTDGYVET 2185
QY 1513 CGSGLÖNR--NYSQÖELLKRVVDY-----EEQÖLESGPHDLFTETYLPROD-----L 1558
D 2186 LKGFTYQSKLTFFSSASITVONADTYVNFAPKHEKESL--NAIRDKELLYLFHÖNSDISIV 2243
QY 1559 EGTPLYESGISLFPDDPESDPEDEAPASAR--VGNIPSTSAKLVQÖLVAESAQSPAA 1617
D 2244 EGG--VQNNMLALY-----DKLINEKREMDLYNINISFTKLQÖE----- 2280
QY 1618 HTTDTAG-----YNAEVSREKPELTAISTERVKNR-----SNVYSGL--TPEEMLY 1665
D 2281 HSTDVFKPMIELHKGNETNKNKSLLEKELKLSVNDHMSMEAEMLKNGLYTFESYONI 2340
QY 1666 -----YKPAKRNH-----ITLITN-----LITEETHVY 1688
D 2341 NNITYIEAEVKTLEEIDRDYDNTQIVEEHKQPSILIDRNALMDIEIFKKNENNL 2400
QY 1689 MKTDAEFVCERTLKYFLGIAGKRVY---VSYPVYQSIKERKMLNEDHFEVGRGVNGRN 1745
D 2401 MEVNETTI--HRVNDYIEKTT--NKLVOAKTEYQÖLENNKÖNDMDQNFETKVSII--EY 2456
QY 1746 HÖGPKRARES 1755
D 2457 FENVKKKKES 2466

RESULT 13
T34418
hypothetical protein F12F3.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T34418
R:Pulton, B.; Wohlmann, P.
submitted to the EMBL Data Library, July 1998
A:Description: The sequence of C. elegans cosmid F12F3.
A:Reference number: 221521
A:Accession: T34418
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-3488 <FNU>
A:Cross-References: EMBL:U080022; PIDD:AA025885.1; GSPDB:GN000023; CESP:F12F3.3
A:Experimental source: strain Bristol N2; clone F12F3
C:Genetics:
A:Gene: CESP:F12F3.3
A:Map position: 5
A:Introns: 281/3; 332/1; 562/3; 600/3; 1866/3; 1944/3; 3393/1

Query Match 3.5%; Score 339.5; DB 2; Length 3488;
Best local similarity 17.8%; Pred. No. 7.8e-07;
Matches 328; Conservative 302; Mismatches 733; Indels 475; Gaps 68;

QY 50 KILNÖKQBPQÖPLCKNDITKRSLOESTRFSQ--VEELIKITICAFQÖDGTGLEANSYNFA 108
D 367 KIIKKKEP-----ESQVYTAPEQÖKISEVDQSVAFTEVGAKKKPDAEKPTDLSKA 420
QY 109 KKENNSPEHLKDEVSIISQMGYRNKAKRLLOSBEPNPSLOTSISVÖLSNGLYRTIYRTK 168
D 421 KDSKSKSKSDPEASTE-----KSTTEKPTNDKTS-----KRSAEKRTYVPK 463

```

```

OY 169 ORI--OPQKTSVYIELGSDSEEDTVNKATYCSVDQELLQITPOG--TRDEILSDAKKA 224
D 464 KEYTGKLEKAKKVEEDKADKASQSPSSKESPPIDGKKKNOI -FKALFIPEIBISRGDPDS 522
OY 225 ACEBSETDVTNTEHHOBSNNDLTETEKRAERHPEKYOGSSVSNLHVEPCGTTHASSIQ. 284
D 523 TMH-SEFINITTTITGRGSDADAKTP-----LVEPIASV----- 555
OY 285 HENSILLTLTDRMNVKAEFC--NKSKOP--GLARSOHNRWAGSKETCNDKRT---PSTE- 337
D 556 ---SMKFTLVESAKKAEFSFKRSETPPDKSRKKEGLPAPKSEKKDVTAEKOSTEA 612
OY 338 ---KKVDLADPLCERKENK-----OKLPCSENPREDTVDYPMITLNSIQVNE 384
D 613 LIESKKKEVEDKISDQPSDKKSEVYVPEKAGETTKDVSEIEVYKKTITKTE 672
OY 385 WFSRDELLGSDSDHDESESNKAVDVLNEVDEYSGSSSEKIDLLASDPHEALICKS 444
D 673 ---KSDSSISQKSNVLPKPADP----- 690
OY 445 ERYHKSVSSEINDFIKKTYRRKKASLPNLSHTENLITAFVTEOIIERPLTNKLR 504
D 691 ---DKSDSDVDK-----SKTTEDQTKVATDSKLEKADTTQIETETVVDKSKK 740
OY 505 KRBPSTGLHDEDFIKKADLAVQKTPKEMINQGTQTEONGQVNNITNSGHENKTKGDSION 564
D 741 KVLKKTREKSDFTISOK-----SETPRVY -EPTKPAEBAOKIAEVNKAQKQKEDVNLKR 795
OY 565 E-----KNPPIESL---EKESAKTKAEBIS--SSISNMELE 597
D 796 EAEVAAKKIADERLKTIEAEANIKKTAEEVAAKKQKEDEQLKLETEVYSSKSAAEKLELE 855
OY 598 LNIHNSAPKKNRLRRKSTRHIALELVYSRNLSPNCELOITDSSSSEELIKKKKYNQ 657
D 856 KQAOIKKAAEADVAKKOKELNEKNKLE--AAKKSADKLKLEESAANKKVSSE- 908
OY 658 MPVRHSNLOIMEGKBPATGAKKSNKPNEQTSKR-----HSDPELTLTN-- 704
D 909 ---SVFGEKKTAKGKKTVOYSEPTSKKTTIDTQVAGTAPADETPKKKIITKK 961
OY 705 ---APGSFTKCSNTSELEKFPVNSLPR -EKEEKELETYKVSNNADPKMLSGER 756
D 962 EKSDSSISQKSAVDSEKVSQKQEDDEPTKPAVSETQMTVADSKKQKETEKLDAEL 1021
OY 757 VLOTBRSVSSSTIYPTDYGTOESTILEVSTLCKAKTERP---KCVSQCAAFENPKG 813
D 1022 AAKTKQADEKS-----KLDAQEKIKKVSDEDAARKERELDKLKESELIAT----- 1068
OY 814 LINGCSKDNENDTEGFKYPLGHEVNSRETSIEMESSELDQAYLQNTFKVSKRQSFALFS 873
D 1069 -----KRASADKLKLEBOQAOKKAAEVEAAKKQKQKEDQKLDTEAASKKAAAEKLE 1120
OY 874 NPGNAEBECATFSAHSGSLKQSPKYTFECEQKQENO---GF-----NESNIKPVQTVNI 925
D 1121 LEKQAOIKKRA---AGADAVYKQ---KELEDEKNKLEANKKSAAGKLTIEESAASKQOTEV- 1174
OY 926 TAFEPVVGQDKRPVNAKCSIKGSRFCSSQFRGMETGLITPNKIGLQNPRIPLP 985
D 1175 -----EQAQ-----LDAQTKAR---TAEKOTKLEKDEKSTKESE 1205
OY 986 IKSFVTKCKKNLLENFEEHSMSPERMENENIPTVSTISNNNIRENVEFKAASSNIN 1045
D 1206 SKETVEKPKKVKLKKKTEKSD-----SSISQKS-----ETSKTYAE 1242
OY 1046 EVG---STINEVGSSSINEIGSSDE-----NIOAELGRNNGPYLNLRLGVLQPEVYKQ 1096
D 1243 SAGPSESETQKVADARKQKETEDEKQKLEAETIYAKKSADEKSKLEAASKL-----KK 1294
OY 1097 SLGSGNCKHPEIKKQKQEEVYQTVNDFSPYLLSDNLEQPMGSSHAQSQSETPDDLDD 1156
D 1295 AAEVEAK-----KQEKDEQLKLDTEAASKKAAAEKLELEK -OSHIIKAA-----EV 1341

```

```

OY 1157 GEIKEDTSAENDIKESSAVFSKS -VQKGLSRSPSPFTTHLAQGRAGAKKLESSEEN 1215
D 1342 DAVKQKLEKQRLSEEAATKKADAEKLTLEQKKAAALILE-----IQKQEK 1393
OY 1216 LSSDEELPCFOHLFEGKVNINIPDSQTRHSTVATECLSKTEENLSLKNLSDCSNOVT 1275
D 1394 LAOBSRL-----EDEAKKSAKQKLESEFTSKQTEE---APKESVDEKPKKV 1439
OY 1276 LAKASQ--EHLISEFTKCSALFSQSCSELEDLTANTNTPDPLIGSSKQMRQSESGV 1333
D 1440 LKKTKESDSSISQKSSAKSTYDA-----AETLESQNLVEKTKVQKVEQSPDESISAT 1495
OY 1334 GLSDKELVSDDEERGTLGE-----NNOEDQSMDSNLGEAASGCESESVSEDCSGLSSQS 1389
D 1496 KROPACKTEISKQDDODEKTTTDDCKPRKPESEATPKRVYKKTQKSDSVASDASLA 1555
OY 1390 DI--LTTQQRDTMOHNLKIQOEMAELEAVLEQHSQSPNSYPSIISDSSALEDLRNPQ 1447
D 1556 DVSKLSDVDEKPKKVKLKKTEKSD-----SVISETSSVDTIK -PES 1597
OY 1448 ---STSEKAVLTSQSSSEYPISONPEGLSADKFEVSADSSSKKEPV-----ERSSP 1498
D 1598 VEIPTKAEQMLLNHRSTDSAVESEPKNAHKDDTEXTTDDMTTRKSSAIFSDDEQIS 1657
OY 1499 SKCPSLDDR-----WYHSCGSLQNRN 1521
D 1658 SKTSSSEGRRRRRRTGFAKASDTLALRGDNVEIEAELEADDTVMKVKGRADLNSRC 1717
OY 1522 YPSQBELIKVVDVEEQOLEESGPHDLET-----SYLPRDLEGT 1561
D 1718 HEMSHTEFFRLIIDEVEPTDSG -MEITANCGETSHHTILKEELPVDFVYKYLPRK -TSGK 1775
OY 1562 PYLESGISLSDPP-----ESDPSDRAPESARVNINISSSALKVPOLKVAESA 1611
D 1776 EGQEVITISVTLNHPIDISKVYWLKDGKPLEINKOYS -IDTVGCSVS-LTLRRAKYEDSG 1832
OY 1612 QSPAAHTTDTAGYNAME-----ESVSRKPELT-----ASTERYNRMGMVVS 1655
D 1833 KYKVVGDGVGCSHTLSIQGKPVAKNVSSEKPVTVTKDQDFSLLVAYDSNPAESFSMTVD 1892
OY 1656 GLTPEEFMLVYKFAKRNHHTLNLITTEETHVVKTD 1693
D 1893 GKDL-----PDGSRSDIVVDGKLTGRGVSKTDA 1923

```

RESULT 14

T38077

hypothetical coiled-coil protein - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999

C:Accession: T38077

R:Conor, R.; Churcher, C.M.; Barrell, B.G.; Rejandream, M.A.; Walsh, S.V.

submitted to the EMBL Data Library, April 1996

A:Reference number: Z21767

A:Accession: T38077

A:Status: preliminary; translated from GB/EMBL/DDJ

A:Molecule type: DNA

A:Residues: 1-1957 <CON>

A:Cross-references: EMBL:Z70690; PTDN:CAA94624.1; GSPDB:GN00066; SPDB:SPACIF3.06c

A:Experimental source: strain 972H-; cosmid c1F3

C:Genetics:

A:Gene: SPDB:SPACIF3.06c

A:Map position: 1

Query Match 3.5%; Score 339; DB 2; Length 1957;

Best Local Similarity 18.5%; Pred. No. 3.6e-07;

Matches 354; Conservative 325; Mismatches 690; Indels 548; Gaps 79;

OY 3 LSAIRVEYQVYINA-----MOKILECP-----CLELIKEPVSTKCHIFCKFMKL 51

D 266 VSTLR--QYENSIRAECKTLQELKCAINEDSKLLEELKHVANYSDAIVHDKDLIED 323

OY 52 LNKQ-----KGPSQCLCKNDITKR-----SLQES-TRFSQVLEELKIIACAFOLD 96

```

Db 324 LSTRSEEDNLKSEEDTLKKNKLEKILNRTIGSLKDSRTSNQLEEEVEL----- 376
Qy 97 TGLEVANSYNFAKKKNNSPEH--LKDEVSIIOSMGVRNRAKRLJOSEPNLSQETSLSV 154
Db 377 -----KESNRTIHSLTDABESKLSFEDENKSLKGSIDEYONNLSKKMKV 423
Qy 155 QLSN-----LGTVRTLRKRIORIQPKTSYIILGSD-----SEEDTVNKAT 195
Db 424 QVSSQLEEARSSLAHATGKLAELINSEBDFQNKIKOFEKIEQDLRACLMSSELKESKA 483
Qy 196 YCSVQDELOLTTPGOTRDEISLDAKKAACESETDVNTEHHQPNNDLTTEKRAE 255
Db 484 LIDKDOELNLRREG- IKQOKRVSESTOSLSQSLQDIIN-----EKK----- 525
Qy 256 RHPEYQSSSVNLHYEPGCTNTHASLSLOHENSLLTKDRMVEKAFCNKSQGLAR 315
Db 526 KH-EYE--SOULKEGE---LOTEISNSELSSQL----- 555
Qy 316 SOHNWAGSKETCNDRTPTSTEEKYVDLADPLCE--RKEMNKOKLPCEPNRDEDPWI 373
Db 556 ---STLAKEKAATAVNTNNELESKNSLOT--LCNAQEKLAQSVQKLKENEQ-----FS 605
Qy 374 TLNASTIQVAVNEFSSSDELGSDSDHSESESNAKAVADVLVYN-----EYDEYSGSSE 427
Db 606 SLDTSPFKRLNE---SHOEL---ENNHOITTKOLKDTSSKLOOLQLEBRANQEKESTLDE 659
Qy 428 KIDILA-----SDPHEALICKSERVHKSQVESNIEKIFGKTYRK-----KALPML 474
Db 660 NNDLTLLKLEESKSLIKOEDV--DSLEKNIO--TLKEDIRKEEALRPSKLEAKML 715
Qy 475 SHVENTLIGAVTPEOIQERPLTNKLRKRRTPTSGLHPEDEFIKADLAVQTP--EMI 532
Db 716 REVINDNL--KCHETLEAQRNDLHSLSDAKN--TALLSELTKSSEDEVKRLTAVENL 771
Qy 533 NOGTNOTEDN-----QVWNITNSGHEKTKGDSIONEKNPNPIESLEKESAFKRAEPI 587
Db 772 TQDSKAMQOFTSLVNSYOSISNLHLEHLDHVNMOSONN---TLLESESKLKTDCENL 827
Qy 588 SSS-----ISNMELEINHSKAPKKNRLRKRSTRIHALELVSNLS----- 632
Db 828 TQOQNTLIDNOVKLMMKHVNOESKYSLEKVGKSLDLKLNLSLNAVASINDOILLLO 887
Qy 633 ---PENCLEQIDSCSSSEELKRYKNQMPVHSHRLQD-----MEGREPATGAKNSK 683
Db 888 AELSKNYDLSBESQOLNSGLKSLKLEAKO--LHTEHEELHRLDKLTGKLTIEESSSL 946
Qy 684 PNEQTSKRHDSPTPELKLTNAPGFTKCSNTSELKEFVNPILPREKEE---KLETVK 739
Db 947 GKLLTARQEE---ISNLKEENNSQOAITSVKSLDETLKSSKLEADIEHLKNKVSVE 1003
Qy 740 VSNNADEPKDMLSGEVLQTERSVSSSISLVPGDYVTOGSSISLLEVSTLGAKATEPN 799
Db 1004 VERNNA-----LASNERLMD----- 1018
Qy 800 KCVSOGAFAENPKGLIHGCSKDNNDTEGFKYPLGHEVHNSRETSIEMESELDAQYON 859
Db 1019 -----DLKNNGERIA-----SLQTELEKRAENDD--LQS 1046
Qy 860 TFKV--SKROSALEFSNPGNAEECATTSANHSGLSKKQSPKYTFECEQKEENQOK--NES 915
Db 1047 KLSVSSSEYENLLISSQTN-----KSLLEDKTNOIKYTEKNVOKKILDEK 1090
Qy 916 NIKPQOTVNITAGFVVGOKDPVDNACSIKGSFCLSSQFRGENTGLIPNKHGILQ 975
Db 1091 DORVNELEELTSKYKLEENQIKDELLALRKSK-----KQMDLCA 1133
Qy 976 NPYRIPLLEPIKSFVTKCKKNLLEENFEHSMSPERENGNNIPSTVSTIRNNIRENV 1035
Db 1134 N-----FY-----DCLKESDALE--QLTNKNKLIVLSLEGSNNNEAL 1170
Qy 1036 FKEMAS--SSNINEVSSSTNEVGSSINELGSDENITQALGRRNRGPKLMLRLVGLQPEV 1093

```

```

Db 1171 VEERSDLANRLSDMKKSLSDSDNVISVIRSLDVRNDEL-----DT 1211
Qy 1094 YKOSLPGSNCKKPEIKROEYEVQTV---NTDFSPYLISDMLBOPMCSHASOVCSETP 1150
Db 1212 LKKDKDSLSTQYSEV--QODRDLIDSLKNGCESFPKKYAVS-----LRELCRSE 1259
Qy 1151 DD-----LLDGEIKEDTSFA-----ENDIKESSAV--FSKSVQKGLSNSPSPFT 1194
Db 1260 IDVPVSEIIDDNFVFNAGNSELRLUTVLSLENYLDANQVAFKMEIDNRLTTTDAEFT 1319
Qy 1195 -----HTHLAQYRKG--AKKLSESEENLSDEDEELPCPHHLLFGVNNIPQST 1242
Db 1320 KYVADLEKLQHEHDDMLLQRGDEKALDKSEKNFLRKAEM-----TEINHSLTEE 1369
Qy 1243 RHSTVATEC--LSKNTENLTS---LKNSLNCSSQVILLAKASQEHHLSEETKCSASLFS 1297
Db 1370 GKEETKKEIAELSSRLIEDNQLATNKLKQD-----HINQELRLKEDVYLK 1414
Qy 1298 SOCS---ELEDDLJANTQDPELLIGSSKQMRHQSSESGVGLDKELVSDDEERGTLLEN 1354
Db 1415 EKESLITSLESLSNQROKESSLDAKNELEH-----MLDDTSRKNSLMEK 1461
Qy 1355 NOE--EOSMDSNLGEAASGCESETSVSEDCSGLSOSDILLTTOQRDPMOHLKLOQEMAE 1413
Db 1462 IESINSLLDKSFELASAVEKLGALOK---LHSESLSIMENIKSQLOEAKETQVDEST 1517
Qy 1414 LEAVLEQHQSPNSY-----PSLISD--SQALEDLKN--PEQSTSEKAVLTSOKSSEY 1463
Db 1518 IOE--LDHETTASKNNYEGKLNKDSIIRDLSENIOQLNNLLAEKSAVKRLSTEKESIL 1576
Qy 1464 PISONPEGLSADKFEVSADSYTSKNKEGVERSPPSKPSLDDRWYMHSCGSLONRNP 1523
Db 1577 QENSRLADEYHKSQVESSELGRSKLK-----LASTTEQLAENE 1616
Qy 1524 SOBELIKVYDVEEQOLEEGPHDLTETSYLPRQDLBGTPLYESGIS----- 1569
Db 1617 RUSLTTRMIDLQO-----VKDLSNIKDSLSEDETLRLSLDEVSASLOKCKEIKSMTVE 1670
Qy 1570 -----LFSDDPESDESEDRAPESA-----RVGNISPTSAKVPOLKVAESAQSPAANT 1619
Db 1671 SLQDVLTVSQAARNALEDEBVSYSVKIRRDORCHLSGKLLHLSQLEBQHEFFRAEQ 1750
Qy 1620 TOTAGYNAMEESVSKPELTAETERNVKNRMSVAVSGLTPEEFMLYKF----- 1668
Db 1731 QRMTOGLFKETVKKO-----EKLLKLNLRQEQILPRSSLIVESYIRDIKEITIV 1782
Qy 1669 --ARKHHITLNLITE-----ETTHVYKMTDAFVCEBRLKYFLGIAGKRWVSY 1716
Db 1783 LOERLNGIELSQOLPKGYFGYFKTNRVEMEVLDSFKQVAKLOFL--AGAETIVKF 1837

```

```

RESULT 15
T30835
breast cancer tumor suppressor Brca2 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 21-Jan-2000
C:Accession: T30835
R:Sharan, S.K.; Bradley, A.
Genomics 40, 234-241, 1997
A:Title: Mutine Brca2: sequence, map position, and expression pattern.
A:Reference number: Z20894; MUID:97237041; PMID:9119389
A:Accession: T30835
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-3328 <SHA>
A:Cross-references: EMBL:U6594; MID:q1743859; PID:q1743860; PIDN:AA023702.1
A:Experimental source: strain C57BL/6
C:Genetics:
A:Gene: Brca2
A:Map position: 5
A>Note: expression of Brca2 detected in midgestation embryos and adult testis, thymus
C:Superfamily: breast cancer tumor suppressor Brca2
C:Keywords: tumor suppressor

```

Query Match 3.5%; Score 334.5; DB 2; Length 3328;  
 Best Local Similarity 18.9%; Pred. No. 1.2e+06;  
 Matches 399; Conservative 286; Mismatches 724; Indels 701; Gaps 96;

QY 72 SLDSEFRESQVLELLKIIC-----AFOLDTGLEIYANSTYPAKKNNS---PEHLKD 120  
 DB 466 SLDSEFRESQVLELLKIIC-----AFOLDTGLEIYANSTYPAKKNNS---PEHLKD 120  
 QY 121 EVSIIOSMGYRNKAKRLLOSEPENPS---LOETSLISVOLSNLGTARTLTKRIQIPQKT 176  
 DB 526 ACGLGILTRACQREDISICPSSVDGSMPTTLTDTSATVK--NAGLISITLKNKR----- 577  
 QY 177 SVYIELGSDSSEDTYKATYCSVGDOLLIQTPQGTREIETSLDAAKACEFSETDV--- 233  
 DB 578 -----KFIY-SVSDASISQGGKLLQTHRQLELNL--SAQLENAFAFEVPL 619  
 QY 234 -TTEHHOPSNNDLNTTEKRAERHPEKYGQSSVSLHVEPCGTNTHASLQHNSSL- 290  
 DB 620 FTWVNSGIPDSSD-----KKRCLEPNDE-----EPSLTNSFGTATSKETISYVIA 663  
 QY 291 LTKDRMNEYKAEFCNKSQOPGLARSOHNRWAGSKETC--NDRRTPTKEKKVLDNADPLCE 349  
 DB 664 LISQD-LNDKEALVITEKPOPTTARADPLLCIPETTCENDOKSPVSN----- 711  
 QY 350 RREWNKOLPCSENPREDTDPVITLNSIIOKVNEFSSDELLGSDSHGSESNAY 409  
 DB 712 -----GKEKVLVS-----ACLPASVQULSISFESQENILG---DHNGSTJLKLTP 753  
 QY 410 ADVLDVLANEVDYSGSSEKIDLLASDPHEALICK--SERVHKSVESNIEDKTFGKT--- 464  
 DB 754 SKRLPL-----SKADWVSREK---MCKMEPEKLOCESKKNIE--LSKNILEY 795  
 QY 465 -----YRKASLPLNLSHVTENLIGAFVTEPOIIOGRPLTNKAKRRRPSGLHPDFTK 519  
 DB 796 NEICITISENSKTPGLPPEBENITIEVASSMKSPNQNAKIVIQDKQSP-----FT- 846  
 QY 520 KADLAQKTPENI--NOGTNOTEQONQVANNITNSGHE-----KTKGDSION----- 564  
 DB 847 -SEVAANMNMSEELFPSPGNNFAQVYNNCKKPKDGLSSVELQEDDLSTHQGPPLAKNPMAY 905  
 QY 565 -----EKNPPIESLEK---ESAFKTKA----- 584  
 DB 906 DEVDVDAHAQVYLITKDSLSLAVVHDYTERKSRNNIIOKHGTEDKCFKSSSLNMMKSDGN 965  
 QY 585 -----EPIS-----STSNMELININNSKAPK--NRLRKSSTR----- 618  
 DB 966 SDCSDKWSFELDVLNHNFGSFRITASNKEIKLSEHNKSKKMFEDIEQYPTRLACID 1025  
 QY 619 HIIHLELVYSRNLSPNCTELQIDSCSSSEIEIKKKRYNNQMPYHRSNLOLMEGKEPATGA 678  
 DB 1026 IYVTLPLANQKKLSEHIFDLK-----SVTIVSTQSHNOSVSH----- 1064  
 QY 679 KSNKNNEBOTSKRHSDTPELKLINAPGSGFTKCSVTSELKEFVNDLPREEKEKLETV 738  
 DB 1065 -----EDDTIAPOM--LSSKODFHNNLTTISQAKETTELSTILLEESGQOFET 1110  
 QY 739 K-----VSNAADEKDLMLSGERYLOT--ERSVSSSISLVPGTGYGOESISLLEVST 790  
 DB 1111 QRRKPSHIQONTSEVGNOMV-----VLSTASKEMKOTDHLPLVDPSPVSGQTDHSKQFEGA 1166  
 QY 791 LCKAKTEPNKCVSQAFAENPKGLIHGCSKDNNDTEGFKY--PLGHEVNHSHETSIEME 848  
 DB 1167 -GVKQSEPHLEBDTCN--KNT-----SCLPLNINEMEFGRICSLGTKLSVSE----- 1212  
 QY 849 ESLEDAQYIANTFKYSKRSQSFALFSPGNAEBECA-----TFSAHSGS-----LKQO 895  
 DB 1213 -----ALRKAMKLFSDIENSEEPAVGRGSSSAHSDSVASVEFKIKRP 1257  
 QY 896 SPKVTFE-----CEQKEENO-----GKN-ESNIKPVQ-----T 922  
 DB 1258 NTEKSFDEKSSKCOVTLQNNIEMKTTCTIYVGRNPEKTYIKNTKHEDSYTSQRRNNLENSDGS 1317

QY 923 VNITAGEPVV--GQDKRPVD-NAKCSIKGSRFELSSQFRGNETGLITENKHLLOMPY 978  
 DB 1318 MSTSG-PVYIHKGDSDLPADQSGKC-----PESC--TOYARENNQIKENISIDLUCLEI 1369  
 QY 979 RIRPLFPISFVATCKKNLLENEFEHNSPBEREMENEPSTVSTISNNINRENVFK 1038  
 DB 1370 MKAETCMKSSDKKOLPSPDKMERNIKEFNIS-----FOYASGKNR--VSK 1414  
 QY 1039 ASSNNINEVSSSTNENEGSSINELGSDPENIOAELGRNRGPKLANMLRGVLOPEVYKQSL 1098  
 DB 1415 SLKSVNIIRRETDL-----TVISLNSKILHGIKDKMHTSCHKAISTIKKVEDHR 1469  
 QY 1099 PGSNCKH-PEIKQOEVEEVQVNTDPSPLIS-----DNLEQPMGSSHASQ 1144  
 DB 1470 PIVTVSQLPAQOQPEVE-----IESRKEPILLFPHNASGKKVIMQESLJK-----YKN 1518  
 QY 1145 VCSETPDOLLDCETIEDTSFAEND--IKSSAVFSKSVQKGLSNPSFFTHILAQGY 1202  
 DB 1519 LCGETO-----YVRKTAFSOGSKPLDKSKKELTVAEKIEVT----- 1556  
 QY 1203 RRGAKKLESSEENLSSDEELPCFOHLLFQKVNINPSOSRHSVATE----- 1250  
 DB 1557 -----ASKCEQONVSKETEMLPQONTHMTQWENTLTSNGTSKVDENIENVEKNPRW 1613  
 QY 1251 CLSKN-----TEENILS--LKNSLNDSNOYILAK----- 1278  
 DB 1614 CIGQSSPYTEDSALAVYTEDSKTCVRESSELSKGRKMLAEQDGLGTRNTIKIECVKEH 1673  
 QY 1279 -----ASQEH-----LSEETKCSASLSSQCELEDL----- 1306  
 DB 1674 TEDFAGNASYEHRLVIIREIDTNHVSNOVSTLLSDPNVCHSYLSOSSSCCHDDMHND 1733  
 QY 1307 -----TAN-----TNTQ----- 1313  
 DB 1734 GFPLAKNIDSDVPPDMKNMAGNTISPRVATKERNLHPQNTINEXCYOKLETNTNSPANKD 1793  
 QY 1314 ---DPLIGSSK-----QMRHOSQGVGLSDRELVSDDERGTGLENNQEDQSD 1362  
 DB 1794 VALDPILLSRNCQKVSIVETITAHQETE-----RTKEIYDNCYKIV-----EQNRQ 1841  
 QY 1363 SNLGEAASGCESTSVSEDCSGLSQSDIITTOORDT-----MOHNLIKLOEAMAL 1414  
 DB 1842 SKPDTCOTSCHKVLDSKDFICPSSSGDVCINRSKDSFCILNEDIOIHN-----QSMS- 1894  
 QY 1415 EAVLEHGQSPNSYSIISDSSALDLRNPQ-----STSEKAVLTSQKSS 1462  
 DB 1895 -----GLKKAATVPVGLTWTDSKSTIREPPQAAHPSTRYGIPTASGKAIVDSASLE 1947  
 QY 1463 YPISQNPBGISADKFEVSADSSSTSKNKEP--GVERSSPSKCPSLDRMYHNSGSGSLQ-- 1518  
 DB 1948 -KARQVFSMDGAKOLSSNVSLGNEKPRHSAKRENS-----VHSTQVLSLP 1996  
 QY 1519 -----NRNTP-----SOELIKVVD-----EEOQLEESG-PHDLTE 1549  
 DB 1997 KPLPGVWNSVSFSGFTAGKLVTVSESAHLKKGKMLEEEDLIRTEHTLTHASPIPEDVSK 2056  
 QY 1550 TSVLPRQDLEGPY--LESGLSLESDDESPSPDRAPESARVGNINIPSSALKVPOLKVA 1608  
 DB 2057 ILPOCAELRTPEYPAVNSKLOKTYNDKSSILPSNYK--ESQSSCNTOSIEVSLDLSOMERN 2114  
 QY 1609 ESASQSPAANTTDTAGYNAMESVSREKPELTASTERVYKRMK-----VSGLTPPEE 1662  
 DB 2115 QDTQL-----VLGTVSHSKANLGLKEDQTLPRKNIKVTKDEMKTEFSDVAVKTN 2161  
 QY 1663 MLVYKFAKRNHITLNLITEETHVYMK-----TDAL-----PYCERTLKYFLCITAG 1710  
 DB 2162 VGEIYKSESENYFEFEAV--ESAKAMEDEDELDSQTHAKSLFPCQNETLFLNSRTK 2219  
 QY 1711 KMWVSFVWTSQIKKEKMLNEND--FEYRGDVVNGNHHGPKRARS---ODKIF---R 1762  
 DB 2220 RGVATVDAAGQPIPKSLINEFDRIIESKGSILT-----PSKSTPDGTGVKDSLFTHHM 2273  
 QY 1763 GLEICCYGPF 1772

Fri Jun 27 10:47:55 2003

us-09-734-672-4.rpr

Page 18

Db 2274 SLEPYTCGPF 2283

Search completed: June 27, 2003, 10:24:30  
Job time : 83 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 27, 2003, 10:03:36 ; Search time 31 Seconds

(without alignments)  
2492.594 Million cell updates/sec

Title: US-09-734-672-4

Sequence: 1 MDLSALRVEQNVINAMOK.....LYQCELDYTLIPQIPSHY 1863

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9642	99.9	1863	BRCL_HUMAN	P38398 homo sapien
2	6923	71.7	1878	BRCL_CANFA	Q95153 canis fam1
3	4817	49.9	1812	BRCL_MOUSE	P48754 mus musculu
4	357.5	3.7	3329	BRCL_MOUSE	P97929 mus musculu
5	340	3.5	2869	RBPI_PLAVB	Q00798 plasmodium
6	339	3.5	1957	YD86_SCHPO	Q10411 schizosacch
7	327	3.4	1875	MLP1_YEAST	Q02455 saccharomyc
8	327	3.4	3418	BRCL_HUMAN	P51587 homo sapien
9	325	3.4	3911	AKA9_HUMAN	P49454 homo sapien
10	323.5	3.4	3210	CENF_HUMAN	Q03661 saccharomyc
11	322	3.3	1658	YME7_YEAST	Q12955 homo sapien
12	321.5	3.3	4377	ANK3_HUMAN	Q12955 homo sapien
13	319.5	3.3	2492	ATRX_HUMAN	P46100 homo sapien
14	317.5	3.3	1972	P531_HUMAN	Q12888 homo sapien
15	316	3.3	1727	ALML1_SCHPO	Q94K55 schizosacch
16	313.5	3.2	1679	Y109_YEAST	P40457 saccharomyc
17	312	3.2	3924	ANK2_HUMAN	Q01484 homo sapien
18	310.5	3.2	2845	APC_MOUSE	Q61315 mus musculu
19	308.5	3.2	2748	NUM1_YEAST	Q00402 saccharomyc
20	308	3.2	2230	GOG4_HUMAN	Q13439 homo sapien
21	307.5	3.2	1790	USO1_YEAST	P25386 saccharomyc
22	304.5	3.2	1744	TANA_XENLA	Q01550 xenopus lae
23	303.5	3.1	1539	Y373_HUMAN	Q15078 homo sapien
24	301.5	3.1	2319	AKA6_HUMAN	Q13023 homo sapien
25	301	3.1	2464	MAPB_MOUSE	P14873 mus musculu
26	299.5	3.1	2843	APC_HUMAN	P25054 homo sapien
27	296.5	3.1	1940	MYH3_HUMAN	P11055 homo sapien
28	294.5	3.1	2349	TPR_HUMAN	P12270 homo sapien
29	293	3.0	2468	MAPB_HUMAN	P46821 homo sapien
30	290	3.0	2663	CENB_HUMAN	Q02224 homo sapien
31	290	3.0	3562	PGCY_CHICK	Q90553 gallus gall
32	288.5	3.0	1928	MYSL_YEAST	P08964 saccharomyc
33	288	3.0	2476	ATRX_MOUSE	Q61687 mus musculu

34	286.5	3.0	1781	1	AKAC_HUMAN	Q02952 homo sapien
35	285.5	3.0	2842	1	APC_RAT	P70478 rattus norv
36	284	2.9	3256	1	KIF6_HUMAN	P46013 homo sapien
37	282.5	2.9	1940	1	MYH3_RAT	P12847 rattus norv
38	282	2.9	1233	1	YF16_YEAST	P43597 saccharomyc
39	281.5	2.9	1934	1	MYH7_MESAU	P13540 mesocricetu
40	280.5	2.9	1935	1	MYH7_HUMAN	P12883 homo sapien
41	280	2.9	1253	1	DSPF_HUMAN	Q9NZ44 homo sapien
42	278.5	2.9	3358	1	PGCY_MOUSE	Q62059 mus musculu
43	276.5	2.9	1935	1	MYH7_PIG	P79293 sus scrofa
44	276.5	2.9	2459	1	MAPB_RAT	P15205 rattus norv
45	275.5	2.9	1433	1	REST_CHICK	Q42184 gallus gall

## ALIGNMENTS

RESULT 1  
BRCL\_HUMAN STANDARD; PRT; 1863 AA.  
ID BRCL\_HUMAN  
AC P38398;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Breast cancer type 1 susceptibility protein.  
GN BRCL.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiindae; Homo.  
OX NCBI\_Taxid=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND VARIANT ARG-1775.  
RX MEDLINE=95025896; PubMed=7545954;  
RA Miki Y., Swensen J., Shattuck-Eidens D., Futreal P.A., Harshman K.,  
RA Tavtigian S., Liu Q., Cochran C., Bennett L.M., Ding W., Bell R.,  
RA Rosenthal J., Hussey C., Tran T., McClure M., Frye C., Hatlier T.,  
RA Phelps R., Haugen-Strano A., Katcher H., Yakumo K., Gholami Z.,  
RA Shaffer D., Stone S., Bayer S., Wray C., Bogden R., Dayanath P.,  
RA Ward J., Tonlin P., Narod S., Bristow P.K., Norris F.H., Helvering L.,  
RA Morrison P., Rostleek P., Lai M., Barrett J.C., Lewis C., Neuhausen S.,  
RA Cannon-Albright L., Godliger D., Wiseman R., Kamb A., Skolnick M.H.,  
RT "A strong candidate for the breast and ovarian cancer susceptibility  
RT gene BRCL.";  
RL Science 266:66-71(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97092865; PubMed=8938427;  
RA Smith T.M., Lee M.K., Szabo C.I., Jerome N., McEuen M., Taylor M.,  
RA Hood L., King M.-C.;  
RT "Complete genomic sequence and analysis of 117 kb of human DNA  
RT containing the gene BRCL.";  
RL Genome Res. 6:1029-1049(1996).  
RN [3]  
RP IDENTIFICATION OF BRCL AS MEMBER OF BASC.  
RX MEDLINE=20245492; PubMed=10783165;  
RA Wang Y., Cortez D., Vazdi P., Neff N., Elledge S.J., Qin J.;  
RT "BASC, a super complex of BRCL-associated proteins involved in the  
RT recognition and repair of aberrant DNA structures.";  
RL Genes Dev. 14:927-939(2000).  
RN [4]  
RP REVIEW ON VARIANTS.  
RX MEDLINE=96400954; PubMed=8807330;  
RA Couch F.J., Weber B.L.;  
RT "Mutations and polymorphisms in the familial early-onset breast  
RT cancer (BRCL) gene. Breast Cancer Information Core.";  
RL Hum. Mutat. 8:8-18(1996).  
RN [5]  
RP VARIANTS LEU-1637; G10-1708 AND ARG-1775.  
RX MEDLINE=95025878; PubMed=7939630;  
RA Futreal P.A., Liu Q., Shattuck-Eidens D., Cochran C., Harshman K.,  
RA Tavtigian S., Bennett L.M., Haugen-Strano A., Swensen J., Miki Y.,  
RA Eddington K., McClure M., Frye C., Weaver-Felhaus J., Ding W.,  
RA Gholami Z., Soederkvist P., Terry L., Jhanwar S., Berchuk A.,

RA Iglehart J.D., Marks J., Ballinger D.G., Barrett J.C., Skolnick M.H.,  
RA Kamb A., Wiseman R.;  
RT "BRCA1 mutations in primary breast and ovarian carcinomas";  
RL Science 266:120-122(1994).  
RN [6]  
RP VARIANTS BC GLY-64, AND VARIANTS ALA-772; ASN-1040 AND GLY-1443.  
RX MEDLINE=95201806; PubMed=7894491;  
RA Castilla L.H., Couch F.J., Erdos M.R., Hoskins K.F., Calzone K.,  
RA Gabler J.E., Boyd J., Lubin M.B., Deshano M.L., Brody L.C.,  
RA Collins J.S., Weber B.L.;  
RT "Mutations in the BRCA1 gene in families with early-onset breast and  
RT ovarian cancer.";  
RL Nat. Genet. 8:387-391(1994).  
RN [17]  
RP VARIANT BC G-61, AND VARIANTS R-356; G-1038; N-1040; R-1183 AND  
RP G-1613.  
RX MEDLINE=95201808; PubMed=7894493;  
RA Friedman L.S., Ostermeyer E.A., Szabo C.I., Dowd P., Lynch E.D.,  
RA Rowell S.E., King M.-C.;  
RT "Confirmation of BRCA1 by analysis of germline mutations linked to  
RT breast and ovarian cancer in ten families.";  
RL Nat. Genet. 8:399-404(1994).  
RN [18]  
RP VARIANT BC GLY-61.  
RX MEDLINE=96108965; PubMed=8554067;  
RA Serova O., Montagna M., Torchard D., Narod S.A., Tonin P., Sylia B.,  
RA Lynch H.T., Feunteun J., Lenoir G.M.;  
RT "A high incidence of BRCA1 mutations in 20 breast-ovarian cancer  
RT families";  
RL Am. J. Hum. Genet. 58:42-51(1996).  
RN [19]  
RP VARIANT BOC TRP-841.  
RX MEDLINE=97123469; PubMed=8968716;  
RA Baker D.F., Almeida E.F.A., Casey G., Fain P.R., Liao S.-Y.,  
RA Masunaka I., Noble B., Kurosaki T., Anton-Culver H.;  
RT "BRCA1 R641W: a strong candidate for a common mutation with moderate  
RT phenotype";  
RL Genet. Epidemiol. 13:595-604(1996).  
RN [10]  
RP VARIANTS BC AND BOC.  
RX MEDLINE=96372821; PubMed=8776600;  
RA Durocher F., Shattuck-Eidens D., McClure M., Labrie F.,  
RA Skolnick M.H., Colagar D.E., Simard J.;  
RT "Comparison of BRCA1 polymorphisms, rare sequence variants and/or  
RT missense mutations in unaffected and breast/ovarian cancer  
RT populations";  
RL Hum. Mol. Genet. 5:835-842(1996).  
RN [11]  
RP VARIANTS BC MET-271 AND SER-1150.  
RX MEDLINE=96303704; PubMed=8723683;  
RA Katagiri T., Emi M., Ito I., Kobayashi K., Yoshimoto M., Iwase T.,  
RA Kasumi F., Miki Y., Skolnick M.H., Nakamura Y.;  
RT "Mutations in the BRCA1 gene in Japanese breast cancer patients";  
RL Hum. Mutat. 7:334-339(1996).  
RN [12]  
RP VARIANT BC GLY-61, AND VARIANTS ARG-239; TRP-841 AND ILE-1512.  
RX MEDLINE=98430998; PubMed=9760198;  
RA Dong J., Chang-Claude J., Wu Y., Schumacher V., Debatin I., Tonin P.,  
RA Royer-Pokora B.;  
RT "A high proportion of mutations in the BRCA1 gene in German  
RT breast/ovarian cancer families with clustering of mutations in the 3'  
RT third of the gene";  
RL Hum. Genet. 103:154-161(1998).  
RN [13]  
RP VARIANT BC GLY-64, AND VARIANTS.  
RX MEDLINE=98141685; PubMed=9482581;  
RA Andersen T.T., Eiken H.G., Couch F., Kaada G., Skrede M., Johnsen H.,  
RA Aloysius T.A., Tveit K.M., Tjanebaerg L., Doerum A., Moeller P.,  
RA Weber B.L., Boerresen-Dale A.-L.;  
RT "Constant denaturant gel electrophoresis (CDGE) in BRCA1 mutation  
RT screening";  
RL Hum. Mutat. 11:166-174(1998).  
RN [14]

RP VARIANTS BC S-22; L-461; D-465; V-552; S-892; D-960; I-1025 AND  
RP A-1047.  
RX MEDLINE=98272917; PubMed=9609997;  
RA Katagiri T., Kasumi F., Yoshimoto M., Nomizu T., Asaiishi K., Abe R.,  
RA Tsuchiya A., Sugano M., Takai S., Yoneda M., Fukutomi T., Nanda K.,  
RA Makita M., Okazaki H., Hirata K., Okazaki M., Furutsuma Y.,  
RA Morishita Y., Iino Y., Karino T., Ayabe H., Hara S., Kajiwara T.,  
RA Houga S., Shimizu T., Toda M., Yamazaki Y., Uchida T., Kunitomo K.,  
RA Sonoo H., Kurebayashi J.-I., Shimotsu K., Nakamura Y., Miki Y.;  
RT "High proportion of missense mutations of the BRCA1 and BRCA2 genes in  
RT Japanese breast cancer families";  
RL J. Hum. Genet. 43:42-48(1998).  
RN [15]  
RP VARIANT OC ARG-1749.  
RX MEDLINE=20455732; PubMed=10486320;  
RA Gayther S.A., Russell P., Harrington P., Antoniou A.C., Easton D.F.,  
RA Ponder B.A.J.;  
RT "The contribution of germline BRCA1 and BRCA2 mutations to familial  
RT ovarian cancer: no evidence for other ovarian cancer-susceptibility  
RT genes";  
RL Am. J. Hum. Genet. 65:1021-1029(1999).  
RN [16]  
RP VARIANT BC S-346, AND VARIANTS L-871; G-1038; R-1183 AND G-1613.  
RX MEDLINE=99254821; PubMed=10323242;  
RA Li S.S.-L., Tseng H.-M., Yang T.-P., Liu C.-H., Teng S.-J.,  
RA Huang H.-W., Chen L.-M., Xao H.-W., Chen J.H., Tseng J.-N., Chen A.,  
RA Hou M.-F., Huang T.-J., Chang H.-T., Mok K.-T., Tsai J.-H.;  
RT "Molecular characterization of germline mutations in the BRCA1 and  
RT BRCA2 genes from breast cancer families in Taiwan";  
RL Hum. Genet. 104:201-204(1999).  
RN [17]  
RP VARIANTS OC, AND VARIANTS.  
RX MEDLINE=99214030; PubMed=10196379;  
RA Janzic S.A., Zlogas A., Krumroy L.M., Krasner M., Plummer S.J.,  
RA Cohen P., Gilda M., Barker D., Haile R., Casey G., Anton-Culver H.;  
RT "Germline BRCA1 alterations in a population-based series of ovarian  
RT cancer cases";  
RL Hum. Mol. Genet. 8:889-897(1999).  
RN [1]  
CC -i- FUNCTION: NOT KNOWN, MAY REGULATE GENE EXPRESSION. INVOLVED IN  
CC TRANSCRIPTIONAL REGULATION OF P21 IN RESPONSE TO DNA DAMAGE.  
CC -i- SUBUNIT: PART OF THE BRCA1-ASSOCIATED GENOME SURVEILLANCE COMPLEX  
CC (BASC), WHICH CONTAINS BRCA1, MSH2, MSH6, MLH1, ATM, BLM, PMS2 AND  
CC THE RAD50-MRE11-NBS1 PROTEIN COMPLEX. THIS ASSOCIATION COULD BE A  
CC DYNAMIC PROCESS CHANGING THROUGHOUT THE CELL CYCLE AND WITHIN  
CC SUBNUCLEAR DOMAINS. CTP1 INTERACTS SPECIFICALLY WITH THE BRC1  
CC DOMAINS.  
CC -i- SUBCELLULAR LOCATION: Nuclear.  
CC -i- POLYMORPHISM: THERE IS EVIDENCE THAT THE PRESENCE OF THE RARE FORM  
CC OF Q356R AND L871P POLYMORPHISMS MAY BE ASSOCIATED WITH AN  
CC INCREASED RISK FOR DEVELOPING OVARIAN CANCER.  
CC -i- DISEASE: BREAST CANCER (BC) IS AN EXTREMELY COMMON MALIGNANCY,  
CC AFFECTING ONE IN EIGHT WOMEN DURING THEIR LIFETIME. A POSITIVE  
CC FAMILY HISTORY HAS BEEN IDENTIFIED AS MAJOR CONTRIBUTOR TO RISK OF  
CC DEVELOPMENT OF THE DISEASE, AND THIS LINK IS STRIKING FOR EARLY-  
CC ONSET BREAST CANCER. MUTATIONS IN BRCA1 ARE THOUGHT TO BE  
CC RESPONSIBLE FOR 45% OF INHERITED BREAST CANCER AND MORE THAN 80%  
CC OF INHERITED BREAST AND OVARIAN CANCER (BOC). MOREOVER, BRCA1  
CC CARRIERS HAVE A 4-FOLD INCREASED RISK OF COLON CANCER, WHEREAS  
CC MALE CARRIERS FACE A 3-FOLD INCREASED RISK OF PROSTATE CANCER.  
CC -i- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
CC -i- SIMILARITY: CONTAINS 2 BRC1 DOMAINS.  
CC -i- DATABASE: NAME=holmcbase; NOTE=BRCA1 entry;  
CC WWW="http://biominformatics.weizmann.ac.il/holmcbase/entries/brca1.htm"  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL: U14680; AAA73985.1; -



Query Match 99.9%; Score 9642; DB 1; Length 1863;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1862; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 MDLSALFVEEYOVYINAMOKILEPCICLLEKEPVSTKCDHIFCKFCMLKLNOKKPSQ 60
DB 1 MDLSALFVEEYOVYINAMOKILEPCICLLEKEPVSTKCDHIFCKFCMLKLNOKKPSQ 60
QY 61 CPLCKNDITKRSLOESRFQSLVEELIKTCAPOLDLGLYANSYNRAKKENSPBLKD 120
DB 61 CPLCKNDITKRSLOESRFQSLVEELIKTCAPOLDLGLYANSYNRAKKENSPBLKD 120
QY 121 EVSLTIOGMYRNARAKRLJOSEPNLSLOETSLVSQSLNGLTVPLRLTKORLOPKTSYVI 180
DB 121 EVSLTIOGMYRNARAKRLJOSEPNLSLOETSLVSQSLNGLTVPLRLTKORLOPKTSYVI 180
QY 181 ELGSDSSEDTVNAKYCVSGDDELQITPOGTRDEISLDSAKKAACEFSETDVTNTEHQ 240
DB 181 ELGSDSSEDTVNAKYCVSGDDELQITPOGTRDEISLDSAKKAACEFSETDVTNTEHQ 240
QY 241 PSNNDLNTTEKRAERPEKRYOGSSVSNLHVEPCGTTHASSLOHENSLLITKDRANVE 300
DB 241 PSNNDLNTTEKRAERPEKRYOGSSVSNLHVEPCGTTHASSLOHENSLLITKDRANVE 300
QY 301 KAERCNKSQOPGLARSOHNMAKSKETCNDRTPTTEKRYDLNADPLCEKKNKOKLPC 360
DB 301 KAERCNKSQOPGLARSOHNMAKSKETCNDRTPTTEKRYDLNADPLCEKKNKOKLPC 360
QY 361 SENPRDTEVPWITLNSIOKVNEWFSRSDDELLGSDSDSHDESESNAKVADVLDVLEVD 420
DB 361 SENPRDTEVPWITLNSIOKVNEWFSRSDDELLGSDSDSHDESESNAKVADVLDVLEVD 420
QY 421 EYSGSSKIDLLASDPHEALICKSERVHKSVESENTEDKTFGKTYRKAALPLNLSHTEN 480
DB 421 EYSGSSKIDLLASDPHEALICKSERVHKSVESENTEDKTFGKTYRKAALPLNLSHTEN 480
QY 481 LIIGAFVTEPOIIIOERPLTNKLRKRRTPTSGLHPEDFIKKRADLAVOTPEMINOQTOTE 540
DB 481 LIIGAFVTEPOIIIOERPLTNKLRKRRTPTSGLHPEDFIKKRADLAVOTPEMINOQTOTE 540
QY 541 QNQVAMNITSGHENTKGTDSIONEKNPNPTESEKSAKTAADPTSSISINMELNLT 600
DB 541 QNQVAMNITSGHENTKGTDSIONEKNPNPTESEKSAKTAADPTSSISINMELNLT 600
QY 601 HNSKAPKRNLRKRSSTRHIALELVYSRNLSPNCTELOIDSCSSSEETKRRKYNOMPV 660
DB 601 HNSKAPKRNLRKRSSTRHIALELVYSRNLSPNCTELOIDSCSSSEETKRRKYNOMPV 660
QY 661 RHSRNLQLMGKBPATGAKKSNKPNEDQTSKRHSDTPELKLITNAPDSFTKCSNTSELKE 720
DB 661 RHSRNLQLMGKBPATGAKKSNKPNEDQTSKRHSDTPELKLITNAPDSFTKCSNTSELKE 720
QY 721 FVNPSPLEPREKEKLETVKVSNNADPKDMLSGERYLOTERSVSSSISLVPGTDYGTQ 780
DB 721 FVNPSPLEPREKEKLETVKVSNNADPKDMLSGERYLOTERSVSSSISLVPGTDYGTQ 780
QY 781 ESISLLEVTSLGAKTEPNKCVSOGCAFENPKGLIHGCSKDNNDPTEGFEYPLGHEVNHNS 840
DB 781 ESISLLEVTSLGAKTEPNKCVSOGCAFENPKGLIHGCSKDNNDPTEGFEYPLGHEVNHNS 840
QY 841 RETSIEMESELDQYLONTFKVSKRSQFALFSPNGAEEECATFSAHSGSLTKQSPKVT 900
DB 841 RETSIEMESELDQYLONTFKVSKRSQFALFSPNGAEEECATFSAHSGSLTKQSPKVT 900
QY 901 FEEDQKFEENGKKNESNIPQYVNTAGFPVYQOKDPVDNACSTIKGSRFLSSQFRG 960
DB 901 FEEDQKFEENGKKNESNIPQYVNTAGFPVYQOKDPVDNACSTIKGSRFLSSQFRG 960
QY 961 NETGLITPNKHGLONPYRIPPLPKSFVTKCKKNLLEENPEEHSMSERMGNNIP 1020
DB 961 NETGLITPNKHGLONPYRIPPLPKSFVTKCKKNLLEENPEEHSMSERMGNNIP 1020
```

```
QY 1021 STVSTISRRNNIRENVFEKASSNINEVGSSSTNEVGSSINELGSSDENIQAELGNRRPKL 1080
DB 1021 STVSTISRRNNIRENVFEKASSNINEVGSSSTNEVGSSINELGSSDENIQAELGNRRPKL 1080
QY 1081 NAMRLGLVLOPEYVKOSLPGSNCKHPEITKKOEVEEVQYVNTDPSPLISDNLEQPMGSS 1140
DB 1081 NAMRLGLVLOPEYVKOSLPGSNCKHPEITKKOEVEEVQYVNTDPSPLISDNLEQPMGSS 1140
QY 1141 HASOVSEPPDDLLDDEIEKEDTSFAENDIKESAVSKSVQGEISRSPPFHTHLAQ 1200
DB 1141 HASOVSEPPDDLLDDEIEKEDTSFAENDIKESAVSKSVQGEISRSPPFHTHLAQ 1200
QY 1201 GYRRGAKKLESSEENLSEDEBELPCFQHLIFGKKNIPISOSTRSTYATCLSKNTEBNL 1260
DB 1201 GYRRGAKKLESSEENLSEDEBELPCFQHLIFGKKNIPISOSTRSTYATCLSKNTEBNL 1260
QY 1261 LSLKNSLNDCSNOVYITLAKASQEHHLSEETKCSASLFFSQCSLEEDLTANTNODPFLIGS 1320
DB 1261 LSLKNSLNDCSNOVYITLAKASQEHHLSEETKCSASLFFSQCSLEEDLTANTNODPFLIGS 1320
QY 1321 SKQMRHQSSEOGVGLSDKELVSDDEERGCTGLEENNOEJQMSDNLGAAAGCESYVSE 1380
DB 1321 SKQMRHQSSEOGVGLSDKELVSDDEERGCTGLEENNOEJQMSDNLGAAAGCESYVSE 1380
QY 1381 DCSGLSSQSDILATTOQBDTMOHNLIKLOQMAELEAVLEOHGSOPSYSITSDSSALE 1440
DB 1381 DCSGLSSQSDILATTOQBDTMOHNLIKLOQMAELEAVLEOHGSOPSYSITSDSSALE 1440
QY 1441 DLNRPQSTSEKAVLTSQKSSEYPISONPEGLSADKFEVSADSTSKNKEPGEVRSRSPSK 1500
DB 1441 DLNRPQSTSEKAVLTSQKSSEYPISONPEGLSADKFEVSADSTSKNKEPGEVRSRSPSK 1500
QY 1501 CPSLDDRWYHWSGSLQNNRNPQOEELIKVVDVEEQLESGBHDLTETSYLPRDLEG 1560
DB 1501 CPSLDDRWYHWSGSLQNNRNPQOEELIKVVDVEEQLESGBHDLTETSYLPRDLEG 1560
QY 1561 TPYLESGISLFPDDPESDPEEDRAPESARVGNIPSSSTALKVQOLKVAESAQSPAANHPT 1620
DB 1561 TPYLESGISLFPDDPESDPEEDRAPESARVGNIPSSSTALKVQOLKVAESAQSPAANHPT 1620
QY 1621 DTAGYNAMESVSREKELTASTERYNKRMSVYVSGITPEEFMLVYKFAKHHITLTNLI 1680
DB 1621 DTAGYNAMESVSREKELTASTERYNKRMSVYVSGITPEEFMLVYKFAKHHITLTNLI 1680
QY 1681 TEETTHVVMKTDAEPVCEERTLKYFLGTAGGKWVSVYVWQSIKERMLNEHDFEVAGDV 1740
DB 1681 TEETTHVVMKTDAEPVCEERTLKYFLGTAGGKWVSVYVWQSIKERMLNEHDFEVAGDV 1740
QY 1741 VNGRNHQPKRARESDORRIFRGLEICCYGPTNMPDQLEMMVQOLCGASVRELSSFTL 1800
DB 1741 VNGRNHQPKRARESDORRIFRGLEICCYGPTNMPDQLEMMVQOLCGASVRELSSFTL 1800
QY 1801 GTGVHPITVVQPDAMTEDNGFHAIGOMCEAPVVTRENVLDLSVALYOCQELDTYLLIPOIPH 1860
DB 1801 GTGVHPITVVQPDAMTEDNGFHAIGOMCEAPVVTRENVLDLSVALYOCQELDTYLLIPOIPH 1860
QY 1861 SHY 1863
DB 1861 SHY 1863
```

RESULT 2  
BRC1\_CANFA  
ID BRC1\_CANFA STANDARD; PRT; 1878 AA.  
AC 095153;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Breast cancer type 1 susceptibility protein homolog.  
CN BRC1.  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

Query Match	Similarity	Score	DB	Length
Best Local	74.2%	Pred. No. 2,5e-282		
Matches 1390:	Conservative 165:	Mismatches 301:	Indels 18:	Gaps 13:
Query	1	MDLSALRREBYOVYINAMOKITLCEPICLELKEIPVSTKCHDHPCKFCMKILNOKKPSQ	60	
Db	1	MDLSADRYEEYOVYINAMOKITLCEPICLELKEIPVSTKCHDHPCKFCMKILNOKKPSQ	60	
Query	61	CPCLKNDITKRSLOESTREFSOLVLELLKITAPOLDGCLFVANSYNAKKENNSPEHLKD	120	
Db	61	CPCLKNDITKRSLOESTREFSOLVLELLKITHAFELDTGLQFADSYNFSKKNNSPEHLKE	120	
Query	121	EVSIIOISMGYNRRKAKRLLOESPENPSLOETSLVQLSNLTGVPRLTRKORIOPOKTSVYI	180	
Db	121	EVSIIOISMGYNRRKAKRLROSEPNPRL-ETISLSVQLSNLTGLVBSRLTRKQOIQOPNKSUYI	179	
Query	181	ELGSDSDSDYVNAKATYCSVGDOELLQITPOGTRDEISLSAKKACAFSEETDVNTEHHQ	240	
Db	180	ELGSDSDSDYVNAKASCSVGDD-LEITSQGARAAEASINPAKKAACEFS-CDITINIEHHQ	237	
Query	241	PSNNDLNTTEKRAAERHPEKQSSVSNLHYEPGCTNTHASLSOHENSSLLTTRDRNVE	300	
Db	238	SSNNDLTTEKHAARKHPEKQYGSVSNLHYEPGCTNTHASLSOHENSSLLTTRHRNVE	297	
Query	301	KAFCFNKSKQGLARSOHNRRAGSKETCNDRPSTETKDYLDNDPLCERKRNKOKLPC	360	
Db	298	KALICNNKSKQGLARSOGRRAESKETCNDRPQISTEKKYVYVNDLDCGRKELNKKQKPPH	357	
Query	361	SENPDTEVDVWITLNNSSIQKVNEMFSKSDLLGSDSDSHOGESESNKAVADVLVDLNEVD	420	

Dd	368	SDSPDSDQVWPVITLNSSTRKVENFSSHDELLTSDSHDRCSELTNTEGCAVEPNEVG	417
Qy	421	EYSGSSEKIDLLADSPHEALICSESRVSKSYESNEIDKIFGTYRRKASLPNLSHTEN	480
Dd	418	EYSGSSEKIDLMAOSDPODAFICESRVRHVKPYGQNIEDIKIFGTYRRKASLPKSHTEYV	477
Qy	481	LIGAFVTEPQIIOERPLTNKLKRRRPTSGLHPEDFTKADLA-VOKTPREMIINOCTOT	539
Dd	478	LFIGACALIEPQMOQHPPMKNAEHRKRRTTSSLHPEDFTKKEVLGIVPTEPKLEIGQIOI	537
Qy	540	EONGVMNITNGSHENKTKDGSIOEKRPNPRESIEKSAFPTKAEPTLSSISNMNLELN	599
Dd	538	KRDGVHINTNGPNETEGDVQOEKANPNPESLEKSAFPTKAEPTKEMPSSRISNMNLELN	597
Qy	600	IHNSKAPKKNRLRRKSSSTRHIALLVVSRNLPSPNCTELODSCSSSEELIKKKKNYOM	659
Dd	598	SSSKAPKKNRLRRKSSSARHICALFEVYNRNLPNDHSELQIESSSSEEMKKNOLDOVP	657
Qy	660	VRHSNLIOMEGKEPATGAKKSNKPNEGOTSKRHDSOTPEELKLTNPAGSFTKCSNTSELK	719
Dd	658	VRHNNTLOMDQKEBAGRAKKSSKRGEOINKLASHAPPELTLTIVSGFFANYSSSSKPQ	717
Qy	720	FEVNSLREEREKELETVKYSNNAEDKDLMSERVLQTPRSVSSSISLYPTDVTG	779
Dd	718	FCINGLRRELEIEESRMTQVSDSRDRKEVLVLSGGRLOTERSEVSTISLYLTDVTGT	777
Qy	780	QESILIEVSTIGRKATEPNCVSCGAAFPENKGLIHGCSKRNNDTGCFKPLGHEVNH	839
Dd	778	ODSILLEADTLIRKAKTYVSNQOANICALIENKPEIHGCSKDRNDTGEFVYPLTRKDNH	837
Qy	840	SRETSIEMESELDAQVLONTFEKYSKROSFALSNPNGAEECATFSAHSGSLKQSPKV	899
Dd	838	TOETSIEMESELDTQCLRNMFVKYSKOSFALFEXPRDPEEDCVTCRSGAFKQKPKV	897
Qy	900	TPECQKRENOCKMNSNTRKYQTVNITAGFPYVGOK-DKPYDNACSTIKGSRPLSSQF	958
Dd	898	TLECQKREESQCKSEIRHVQAVHTNAGFSAVSQAKKPGGFAKCSIKGYSRLCLSSQF	957
Qy	959	RGNENGLITPNKHGLQOYRILPILPFIKSFATKCKNLLEENFEHSEMSPEREMGEN	1018
Dd	958	KGKETELLIANHGISQNPYHPLPLPSIRSCYKTLCOENLSBEKEQJMSPERAVGNER	1017
Qy	1019	IPSTVSTISRRNIRENVFKEAASSNINEVGSSTNEVGSSINEIGSSDENITQAEIGRRNG	1077
Dd	1018	VIOSTVSTISQNNIRECASKESVSSNVEVYSTNEVGSSVNEVGSSEENITQAEIGRRNG	1077
Qy	1078	PCLNMLLELVYQPEPVYKQSLPGSKCKRPELIKQ-EYEVYQTVNTEDSPILISNLEOP	1136
Dd	1078	PCLNMLLELVYQPEPVYKQSLPGSKCKRPELIKQ-EYEVYQTVNTEDSPILISNLEOP	1137
Qy	1137	MGSSHASOVCSPTPDLLDDECFIKDTPFAENDIKESSAVFSKSVQKELSRSPPHYT	1196
Dd	1138	MGSSKSSSOVCSETTPDOLLNGDIKGVKVSFAESDILKEKSAVFSKSVQSEFBSRSPSOHT	1197
Qy	1197	HLAOGYRRGAKKLESSEENLSSEDEELPCFOHLLFGKVNINLPQSOTRSTVATECLSKNT	1256
Dd	1198	RLAOGYQGTGKTKLESSEENMSSEBELPCFOHLLFGKXTNMPQSOTSHINAVAAGLSKMT	1257
Qy	1257	EENLISLKNISLDCSNQVILAKASOEHNLSBYTKCSASLESFOSCELEDLTANTNTODPF	1316
Dd	1258	EENLIDSLKNSLDSISNQVPSAKASOEHNLSBEARCSGSLFSSQCSALEDLTVNNTODPF	1317
Qy	1317	LI--GSSKOMHOSGQGVGLSDKFLVS-DOEBRGTYLEENN-DOEOMSDNILEAASGC	1372
Dd	1318	SMEDPTSKQVHRQSENLDV-LNDXELVSDDDDERPGLLEBDSPOEBQSVVDVLGVAAGY	1376
Qy	1373	ESENTSVSEDCSGLSQSDILTTQOQRPDTQOHNLIKLOQMALEAVLEBHGQSPNSYPSI	1432
Dd	1377	ESENTSLSDCGRLSQSDILTTQOQRPDTQOHNLIKLOQMALEAVLEBHGQSPNSYPSL	1436
Qy	1433	ISDSSALEDLRNPEDOSTEKAVALTSQKSEYPISONPEGLISADKFEVADSSTKKNPEG	1492

Db 1437 IADSCSPEDLLNPEONASER-VITSEKSDSPISQNBESLSTDKFOVFLDSTSKNCEPG 1495  
 QY 1493 VERSSPKCPLDDRWYMHSCSGSLQNRNYPQSEELIKVYDVERQQLSESGPHDLTSTY 1552  
 Db 1496 MINSPPQSRLDLTRMYHSCPSKSLQDTPNCPQSEELIKVYDVERQQLSESGPHDLTSTY 1555  
 QY 1553 LPRDLEGTFFYLSGLSLFSDDEPESDEAPASAVGNIPSTSLAKYPOKVAESAQ 1612  
 Db 1556 LSRDLEGARYLSEGLSLFSDDEPESDPSHRASLAVSSMPSTSLAKYPOKVAESAQ 1615  
 QY 1613 SPAAHTTDTAGTAVNAMESVREKPELTASTERYNKMVNSQLAPEEFMLYKFAK 1672  
 Db 1616 STAAVHTASTAGYKNSDSESVIEKPEVISTRGVKNRISWASGLPKEFMLYKFAK 1675  
 QY 1673 HITLTNLTETHTVYVVKTAEPFCELTLYFLGIAGKVVSVFWTOSIKERKMLNEH 1732  
 Db 1676 HISTLTNLTETHTVYVVKTAEPFCELTLYFLGIAGKVVSVFWTOSIKERKMLNEH 1735  
 QY 1733 DFEYRGDVNVRNHOQPKRA-----RESODRKIFRGLEICCGFPTNMPDLEMMVQLC 1787  
 Db 1736 DFEYRGDVNVRNHOQPKRARRESQDRSKIFRGLEICCGFPTNMPDLEMMVQLC 1795  
 QY 1788 GASVYKASLFTLGTGVHPVYVQPDAMTENGHFGMCCEAPVYTRRWVLDVALYQC 1847  
 Db 1796 GASVYKASLFTLGTGVHPVYVQPDAMTENGHFGMCCEAPVYTRRWVLDVALYQC 1855  
 QY 1848 QELDTYLIPOIPHS 1861  
 Db 1856 QELDTYLIPOIPRT 1869

## RESULT 3

BRC1\_MOUSE STANDARD: PRT: 1812 AA.

ID BRC1\_MOUSE  
 AC P48734; Q60957; Q60983;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Breast cancer type 1 susceptibility protein homolog.  
 GN BRCAL.  
 OS Mus musculus (mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxId=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6; TISSUE=Embryo;  
 RX MEDLINE=96177659; PubMed=8634697;  
 RA Abel K.J., Xu J., Yin G.Y., Lyons R.H., Meisler M.H., Weber B.L.;  
 RT "Mouse Brcal: localization sequence analysis and identification of  
 evolutionarily conserved domains";  
 RL Hum. Mol. Genet. 4:2265-2273(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6;  
 RX MEDLINE=96177660; PubMed=8634698;  
 RA Sharan S.K., Wims M., Bradley A.;  
 RT "Murine Brcal: sequence and significance for human missense  
 mutations";  
 RL Hum. Mol. Genet. 4:2275-2278(1995).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129/SVJ;  
 RX MEDLINE=96121367; PubMed=8575748;  
 RA Bennett L.M., Haugen-Striano A., Cochran C., Brownlee H.A.,  
 \*Fiedorek F.T., Jr., Wiseman R.W.;  
 RT "Isolation of the mouse homologue of BRCAL and genetic mapping to  
 mouse chromosome 11";  
 RL Genomics 29:576-581(1995).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129/SVJ; TISSUE=Embryo;  
 RX MEDLINE=96067162; PubMed=7590247;

RA Lane T.F., Deng C., Elson A., Lyu M.S., Kozak C.A., Leder P.;  
 RT "Expression of Brcal is associated with terminal differentiation of  
 ectodermally and mesodermally derived tissues in mice";  
 RL Genes Dev. 9:2712-2722(1995).  
 RN [5]  
 RP SEQUENCE OF 727-1111 FROM N.A.  
 RC STRAIN=C57BL/6; TISSUE=Embryo;  
 RX MEDLINE=96021028; PubMed=7550308;  
 RA Marquis S.T., Rajan J.V., Wynshaw-Boris A., Xu J., Yin G.Y.,  
 RA Abel K.J., Weber B.L., Chodosh L.A.;  
 RT "The developmental pattern of Brcal expression implies a role in  
 differentiation of the breast and other tissues";  
 RL Nat. Genet. 11:17-26(1995).  
 RN [6]  
 RP SEQUENCE OF 789-1250 FROM N.A.  
 RC STRAIN=129/SVJ;  
 RX MEDLINE=96163506; PubMed=8566965;  
 RA Schroeder E., Badger P., Larson D., Erdos M., Wynshaw-Boris A.,  
 RA Red T., Brody L.;  
 RT "The murine homologue of the human breast and ovarian cancer  
 susceptibility gene Brcal maps to mouse chromosome 11D.";  
 RL Hum. Genet. 97:256-259(1996).  
 CC -1- FUNCTION: NOT KNOWN, MAY REGULATE GENE EXPRESSION. INVOLVED IN  
 TRANSCRIPTIONAL REGULATION OF P21 IN RESPONSE TO DNA DAMAGE (BY  
 SIMILARITY).  
 CC -1- SUBUNIT: CIP1 INTERACTS SPECIFICALLY WITH THE BRC1 DOMAINS (BY  
 SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Nuclear (potential).  
 CC -1- TISSUE SPECIFICITY: IN THE EMBRYO, EXPRESSION DECREASES IN OTIC VESICLES AT  
 DAY 9.5. AT DAY 10.5, THIS EXPRESSION DECREASES AND HIGH LEVELS  
 ARE FOUND IN THE NEUROECTODERM. AT DAYS 11-12.5, HIGH LEVELS IN  
 DIFFERENTIATING KERATINOCYTES AND WHISKER PAD PRIMORDIA. AT DAYS  
 14-17, EXPRESSION ALSO OBSERVED IN KIDNEY EPITHELIAL CELLS. IN  
 THE ADULT, HIGHEST LEVELS FOUND IN SPLEEN, THYMUS, LYMPH NODES,  
 EPITHELIAL ORGANS, AND ALVEOLAR AND DUCTAL EPITHELIAL CELLS OF  
 THE MAMMARY GLAND. VERY LOW LEVELS IN BRAIN, KIDNEY, AND SKIN. NO  
 EXPRESSION IN HEART, LIVER OR LUNG.  
 CC -1- DEVELOPMENTAL STAGE: IN THE MAMMARY GLAND, EXPRESSION INCREASES  
 DRAMATICALLY DURING PREGNANCY. LEVELS FALL DURING LACTATION AND  
 INCREASE AGAIN DURING POST-LACTATIONAL REGRESSION OF THE  
 MAMMARY GLAND.  
 CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
 CC -1- SIMILARITY: CONTAINS 2 BRC1 DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U35641; AAB17113.1; -  
 DR EMBL: U31625; AAB17114.1; -  
 DR EMBL: U32446; AAA96393.1; -  
 DR EMBL: U36475; AAC52323.1; -  
 DR EMBL: U35835; AAA95742.1; -  
 DR MGD: MGI:104537; Brcal.  
 DR InterPro: IPR001357; BRC1.  
 DR InterPro: IPR002378; Brc1.  
 DR InterPro: IPR001841; Znf\_finger.  
 DR Pfam: PF00097; zf-C3HC4; 1.  
 DR Pfam: PF00533; BRC1; 2.  
 DR PRINTS: PR00493; BRSTCANCER1.  
 DR SMART: SM00292; BRC1; 2.  
 DR SMART: SM00184; RING; 1.  
 DR PROSITE: PS50172; BRC1; 2.  
 DR PROSITE: PS50089; ZF\_RING\_2; 1.  
 DR PROSITE: PS50089; ZF\_RING\_2; 1.  
 DR Zinc-finger; DNA-binding; Nuclear protein; Anti-oncogene; Repeat.  
 KW ZN\_FING; 24 65  
 FT DOMAIN 1585 1679 BRC1 1.  
 FT DOMAIN 1698 1797 BRC1 2.

ET VARIANT 93 93 F -> L (IN STRAIN 129/SVJ) .  
ET VARIANT 305 305 T -> S (IN STRAIN 129/SVJ) .  
ET VARIANT 319 319 P -> A (IN STRAIN 129/SVJ) .  
ET VARIANT 377 377 Q -> E (IN STRAIN 129/SVJ) .  
ET VARIANT 550 550 K -> Q (IN STRAIN 129/SVJ) .  
ET VARIANT 652 652 A -> P (IN STRAIN 129/SVJ) .  
ET VARIANT 765 765 S -> P (IN STRAIN 129/SVJ) .  
ET VARIANT 917 917 P -> L (IN STRAIN 129/SVJ) .  
ET VARIANT 933 933 C -> S (IN STRAIN 129/SVJ) .  
ET VARIANT 1122 1122 K -> I (IN STRAIN 129/SVJ) .  
ET VARIANT 1206 1206 S -> R (IN STRAIN 129/SVJ) .  
ET VARIANT 1212 1212 RM -> GI (IN STRAIN 129/SVJ) .  
ET VARIANT 1255 1255 S -> R (IN STRAIN 129/SVJ) .  
ET VARIANT 1261 1261 H -> N (IN STRAIN 129/SVJ) .  
ET VARIANT 1264 1264 V -> A (IN STRAIN 129/SVJ) .  
ET VARIANT 1269 1269 P -> A (IN STRAIN 129/SVJ) .  
ET VARIANT 1283 1283 T -> K (IN STRAIN 129/SVJ) .  
ET VARIANT 1337 1337 T -> N (IN STRAIN 129/SVJ) .  
ET VARIANT 1349 1349 P -> T (IN STRAIN 129/SVJ) .  
ET VARIANT 1352 1353 EG -> QR (IN STRAIN 129/SVJ) .  
ET VARIANT 1381 1381 S -> P (IN STRAIN 129/SVJ) .  
ET VARIANT 1390 1390 G -> A (IN STRAIN 129/SVJ) .  
ET VARIANT 1400 1400 V -> D (IN STRAIN 129/SVJ) .  
ET VARIANT 1503 1503 E -> Q (IN STRAIN 129/SVJ) .  
ET VARIANT 1549 1549 V -> A (IN STRAIN 129/SVJ) .  
ET VARIANT 1680 1680 T -> K (IN STRAIN 129/SVJ) .  
ET VARIANT 1712 1712 D -> E (IN STRAIN 129/SVJ) .  
ET VARIANT 1721 1721 D -> E (IN STRAIN 129/SVJ) .  
SQ SEQUENCE 1812 AA: 198669 MM: 2291EA74150B86A CRC64:

Query Match 49.9% Score 4817; DB 1; Length 1812;  
Best Local Similarity 55.5%; Pred. No. 2,9e-194;  
Matches 1035; Conservative 261; Mismatches 498; Indels 70; Gaps 32;

OY 1 MDLSLRVEYONVNNAMOKILECPICILELKEPVSTKCDHIFCKRCMKLNNOKKGPQ 60  
DB 1 MDLSLRVEYONVNNAMOKILECPICILELKEPVSTKCDHIFCKRCMKLNNOKKGPQ 60  
OY 61 CPLCNDITKRSLOESTRFSQVLEELKICAFOLDTGLEVANSYFAKKENNSPHELD 120  
DB 61 CPLCNDITKRSLOESTRFSQVLEELKICAFOLDTGLEVANSYFAKKENNSPHELD 120  
OY 121 EYVSIISQMGYRNRAKRLLOSPEPNSLOETSLVSLOLNTGVTLETKRKRIOPKTSVYI 180  
DB 121 EYVSIISQMGYRNRAKRLLOSPEPNSLOETSLVSLOLNTGVTLETKRKRIOPKTSVYI 180  
OY 121 EASIIOSGYRYNRVRLPQVEBPNGATLKD-SLGVOLSNLGIYVSVYKKNROTQPRKKSXYI 179  
DB 121 EASIIOSGYRYNRVRLPQVEBPNGATLKD-SLGVOLSNLGIYVSVYKKNROTQPRKKSXYI 179  
OY 181 ELGSSSEDTYKATYCSYGOELLQTPQGTREDELISLAKKAACEFSETDVTNFEHNO 240  
DB 181 ELGSSSEDTYKATYCSYGOELLQTPQGTREDELISLAKKAACEFSETDVTNFEHNO 240  
OY 180 ELDSOSSSEETVTKPGDCSYRDELLQTAPOEAGDEBKLHSAEBAACEFSE-GIRNLEHNO 238  
DB 180 ELDSOSSSEETVTKPGDCSYRDELLQTAPOEAGDEBKLHSAEBAACEFSE-GIRNLEHNO 238  
OY 241 PSNNDLNTTEKRAAEHRPEKYOGSSVSNLHVEPCGTNTHASSLOHENSLLLTCKDRMNYE 300  
DB 241 PSNNDLNTTEKRAAEHRPEKYOGSSVSNLHVEPCGTNTHASSLOHENSLLLTCKDRMNYE 300  
OY 239 CS-DLPLNTFNHATERHPEKCCSISINVCPECGTDAAHSSLOPETSSLLLEDRMNE 297  
DB 239 CS-DLPLNTFNHATERHPEKCCSISINVCPECGTDAAHSSLOPETSSLLLEDRMNE 297  
OY 301 KAEFCNKSKOPGLASQHNRAAGSKETCDNRTPSTFEKVDLADPLCERKEKKNKCLPC 360  
DB 301 KAEFCNKSKOPGLASQHNRAAGSKETCDNRTPSTFEKVDLADPLCERKEKKNKCLPC 360  
OY 298 KAEFCNKSKOPGLASQHNRAAGSKETCDNRTPSTFEKVDLADPLCERKEKKNKCLPC 357  
DB 298 KAEFCNKSKOPGLASQHNRAAGSKETCDNRTPSTFEKVDLADPLCERKEKKNKCLPC 357  
OY 361 SENPDTEADVPIITLNNSTIOKYNEMFSRDELLGSDSHDGESESNAAKVADYVDNEVD 420  
DB 361 SENPDTEADVPIITLNNSTIOKYNEMFSRDELLGSDSHDGESESNAAKVADYVDNEVD 420  
OY 358 PENSATITDVPWITLNNSTIOKYNEMFSRDELLGSDSHDGESESNAAKVADYVDNEVD 417  
DB 358 PENSATITDVPWITLNNSTIOKYNEMFSRDELLGSDSHDGESESNAAKVADYVDNEVD 417  
OY 421 EYSGSSEKIDLLASDPHEALICKSERVHSKSVESNIEDKIFGTYRKKASLPNLVHTEN 480  
DB 421 EYSGSSEKIDLLASDPHEALICKSERVHSKSVESNIEDKIFGTYRKKASLPNLVHTEN 480  
OY 418 GGFSSSRKTDVTPRPHHTLCKSGRDPFKPYEDNITSKIFGYSYRKSRPLNHYTE- 476  
DB 418 GGFSSSRKTDVTPRPHHTLCKSGRDPFKPYEDNITSKIFGYSYRKSRPLNHYTE- 476  
OY 481 LIGAFVIEPOITIOERPLTNKLRKRPRPSGLHPDEFIKKADIA-VQKTPREMIINOCT 539  
DB 481 LIGAFVIEPOITIOERPLTNKLRKRPRPSGLHPDEFIKKADIA-VQKTPREMIINOCT 539  
OY 477 -IIGFTIEPOITIOERPLTNKLRKRPRPSGLHPDEFIKKADIA-VQKTPREMIINOCT 533  
DB 477 -IIGFTIEPOITIOERPLTNKLRKRPRPSGLHPDEFIKKADIA-VQKTPREMIINOCT 533  
OY 540 EONGVNMITNSGCHENKTKGDSIONENKPNPIESLEKESAFKTKAPRISSISNMLEJLN 599  
DB 540 EONGVNMITNSGCHENKTKGDSIONENKPNPIESLEKESAFKTKAPRISSISNMLEJLN 599  
OY 534 EPNEDAVSTTSKQENKTAGSNLQEKSAHPRESLKEPEASTAGAKSISNSVSDLEVEJLN 593  
DB 534 EPNEDAVSTTSKQENKTAGSNLQEKSAHPRESLKEPEASTAGAKSISNSVSDLEVEJLN 593

OY 600 IHNSKAPKKNRLRRKSSSTHIALHLELVSNRLSPNCTELQDSCSSSEELIKKKKNOMP 659  
DB 594 VHSKAPKKNRLRRKSSSTHIALHLELVSNRLSPNCTELQDSCSSSEELIKKKKNOMP 652  
OY 660 VHSKAPKKNRLRRKSSSTHIALHLELVSNRLSPNCTELQDSCSSSEELIKKKKNOMP 719  
DB 653 VHSKAPKKNRLRRKSSSTHIALHLELVSNRLSPNCTELQDSCSSSEELIKKKKNOMP 711  
OY 720 EPNEDAVSTTSKQENKTAGSNLQEKSAHPRESLKEPEASTAGAKSISNSVSDLEVEJLN 778  
DB 712 EPNEDAVSTTSKQENKTAGSNLQEKSAHPRESLKEPEASTAGAKSISNSVSDLEVEJLN 770  
OY 779 TQESILLEVSTLGAKEPTEPNKCVSOCAAFENPKGLIHGCKDNKNDDECKRYPLGHEVN 838  
DB 771 TQESILLEVSTLGAKEPTEPNKCVSOCAAFENPKGLIHGCKDNKNDDECKRYPLGHEVN 829  
OY 839 HSRRETSIEHESELEDAQYLONTFVKYSKROSLFSPGNAEBECATFSAHSGSLKQSPK 898  
DB 830 HSRRETSIEHESELEDAQYLONTFVKYSKROSLFSPGNAEBECATFSAHSGSLKQSPK 884  
OY 889 VPFECOKKEENCKNENKPYOTVNTAGPVGOKKOPVONAKCSITKGSRFCLSSQF 958  
DB 885 VPFECOKKEENCKNENKPYOTVNTAGPVGOKKOPVONAKCSITKGSRFCLSSQF 941  
OY 959 RGNETGLTPNKHGLLONFYRIPPLPIKSFYKTKCKKNLLENEFEHSMSPEREMGEN 1018  
DB 942 RGNETGLTPNKHGLLONFYRIPPLPIKSFYKTKCKKNLLENEFEHSMSPEREMGEN 1001  
OY 1019 I-PTVSTISRNRRNREKKAASSNINEVSGSTINEIGSSSDENQALGKRNK 1077  
DB 1002 I-PTVSTISRNRRNREKKAASSNINEVSGSTINEIGSSSDENQALGKRNK 1046  
OY 1078 PKTNAMLRGLVLOPEYKOSLPGSNCKHPEIKKOYEVEYQVNTDFPYLISDLDEPM 1137  
DB 1047 PKTNAMLRGLVLOPEYKOSLPGSNCKHPEIKKOYEVEYQVNTDFPYLISDLDEPM 1101  
OY 1138 GSSHASOVSETPDDLLDDEIKEDTSPFANDIKSSAVFSKVOGELSSPSPFTTH 1197  
DB 1102 GSSHASOVSETPDDLLDDEIKEDTSPFANDIKSSAVFSKVOGELSSPSPFTTH 1160  
OY 1198 LAOGYRGAKKLLESSEENLSEDELPCHLHPKVNINPSOSTRHSATVATECLSKTE 1257  
DB 1161 LAOGYRGAKKLLESSEENLSEDELPCHLHPKVNINPSOSTRHSATVATECLSKTE 1218  
OY 1258 ENLSLKNLNCOSNOVILAKASOEHHLSPEFKCASLPSOSCSELEDTANTNTQDFL 1317  
DB 1219 ENLSLKNLNCOSNOVILAKASOEHHLSPEFKCASLPSOSCSELEDTANTNTQDFL 1278  
OY 1318 IGSSKQMRHOSQGVGLSDKELVSDDEBGTGLENNQEOGSDNLSGEA-ASGCESET 1376  
DB 1279 IGSSKQMRHOSQGVGLSDKELVSDDEBGTGLENNQEOGSDNLSGEA-ASGCESET 1336  
OY 1377 SVSEDCSGSSOSDITLTOQRTQMNHLKLOEWAHELEAVLEQHSOPNSYSIISDS 1436  
DB 1337 SVSEDCSGSSOSDITLTOQRTQMNHLKLOEWAHELEAVLEQHSOPNSYSIISDS 1392  
OY 1437 SALEDLRNEQOSTSEKAVLTSQSSSEYPISQNPELSADKEFVS-ADSTSKNKEPGEV 1495  
DB 1453 SALEDLRNEQOSTSEKAVLTSQSSSEYPISQNPELSADKEFVS-ADSTSKNKEPGEV 1508  
OY 1556 QDEGTPYLESISTLSPD-DEPSDPSEDRAPASAVNGIPSTSLAKYPLQKVAESAOP 1614  
DB 1509 QDEGTPYLESISTLSPD-DEPSDPSEDRAPAPASAVNGIPSTSLAKYPLQKVAESAOP 1564  
OY 1615 AAHTTDTAGYNAMESVSREKPELTASTERYKKNMSVVGSLPDEEMLYTKARKNHI 1674  
DB 1565 AAHTTDTAGYNAMESVSREKPELTASTERYKKNMSVVGSLPDEEMLYTKARKNHI 1617

QY 1675 TLNLTTEETHVYMTDAEFVCEKTLKFLGAGKMWVSIFWVQSIKERKMLNEDH 1734  
 Db 1618 TLTLDAITEETHVITITDAEFVCEKTLKFLGAGKMWVSIFWVQSIKERKMLNEDH 1677  
 QY 1735 EVAGDVVNGRHHGPPRRARESDRKFIRGLICCYGPTFMPPDOLEMMVQLCGASVKE 1794  
 Db 1678 EVAGDVVNGRHHGPPRRARESDRKFIRGLICCYGPTFMPPDOLEMMVQLCGASVKE 1736  
 QY 1795 LSSEFTLGTVHPVIVVOPDAMTEGNGFHAIGMCEAPVYTRRWVLDLSVALYCCOEDTYL 1854  
 Db 1737 LPSLTHDTGALHVLVIQPSAWTEDSNCPDGLGCKARLVMMQVLDLSLSYRCRDADYL 1796  
 QY 1855 IPQI 1858  
 Db 1797 VQNI 1800

RESULT 4  
 BRC2\_MOUSE STANDARD; .PRT; 3329 AA.  
 AC P97929; P97983; 035922;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DE Breast cancer type 2 susceptibility protein.  
 GN BRCA2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129;  
 RX MEDLINE=97217789; PubMed=9063750;  
 RA Connor F., Smith A., Wooster R., Stratton M., Dixon A., Campbell E.,  
 RT Tait T.M., Freeman T., Ashworth A.;  
 RT Cloning, chromosomal mapping and expression pattern of the mouse  
 RT Brca2 gene.";  
 RL Hum. Mol. Genet. 6:291-300(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6;  
 RX MEDLINE=97237041; PubMed=9119389;  
 RA Sharon S.K., Bradley A.;  
 RT Murine Brca2: sequence, map position, and expression pattern.";  
 RL Genomics 40:234-241(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129/SV;  
 RX MEDLINE=97384941; PubMed=9242436;  
 RA McAllister K.A., Haugen-Strano A., Hagevik S., Brownlee H.A.,  
 RT Collins N.K., Futreal P.A., Bennett L.M., Wiseman R.W.;  
 RT Characterization of the rat and mouse homologues of the BRCA2 breast  
 RT cancer susceptibility gene.";  
 RL Cancer Res. 57:3121-3125(1997).  
 RN [4]  
 RP SEQUENCE OF 18-200 FROM N.A.  
 RX MEDLINE=97075121; PubMed=8917547;  
 RA Rajan J.V., Wang M., Marguis S.T., Chodosh L.A.;  
 RT Brca2 is coordinately regulated with Brca1 during proliferation and  
 RT differentiation in mammary epithelial cells.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 93:13078-13083(1996).  
 RN [5]  
 RP SEQUENCE OF 569-625 FROM N.A.  
 RX MEDLINE=97341126; PubMed=9196008;  
 RA McAllister K.A., Ramachandran S., Haugen-Strano A., Fiedorek F.T. Jr.,  
 RA Wiseman R.W.;  
 RT "Genetic mapping of the Brca2 breast cancer susceptibility gene on  
 RT mouse chromosome 5.";  
 RL Mamm. Genome 8:540-541(1997).  
 CC -!- FUNCTION: MAY PARTICIPATE IN A PATHWAY ASSOCIATED WITH THE  
 CC ACTIVATION OF DOUBLE-STRAND BREAK REPAIR AND/OR HOMOLOGOUS  
 CC RECOMBINATION (BY SIMILARITY).

CC -!- SUBUNIT: INTERACTS WITH RAD51 (BY SIMILARITY).  
 CC -!- TISSUE SPECIFICITY: WIDELY EXPRESSED. HIGHEST EXPRESSION IN  
 CC CEREBELLUM, TESTIS, ILEUM, APPENDIX, EPIDIDYMIS, OVARY AND MAMMARY  
 CC GLAND. NO EXPRESSION IN LUNG.  
 CC -!- DEVELOPMENTAL STAGE: IN THE MAMMARY GLAND, EXPRESSION INCREASES  
 CC DRASTICALLY DURING PREGNANCY.  
 CC -!- SIMILARITY: CONTAINS 7 BRCA2 REPEATS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

CC EMBL: U82270; AAB48306.1; -;  
 CC EMBL: U72947; AAB40720.1; -;  
 CC EMBL: U65594; AAC23702.1; -;  
 CC EMBL: U89652; AAB71377.1; -;  
 CC EMBL: U89503; AAC53276.1; -;  
 CC MGD: MGI:109337; Brca2  
 CC InterPro: IPR002093; BRCA2\_repeat.  
 CC Pfam: PF00634; BRCA2; 7.  
 CC PROSITE: PS50138; BRCA2\_REPEAT; 6.  
 CC KW Polymorphism; Repeat.  
 CC FT REPEAT 981 1015  
 CC FT REPEAT 1192 1226  
 CC FT REPEAT 1394 1428  
 CC FT REPEAT 1491 1525  
 CC FT REPEAT 1623 1657  
 CC FT REPEAT 1924 1958  
 CC FT REPEAT 2004 2038  
 CC FT REPEAT 44 44  
 CC FT VARIANT 340 340  
 CC FT VARIANT 377 377  
 CC FT VARIANT 407 407  
 CC FT VARIANT 661 661  
 CC FT VARIANT 739 739  
 CC FT VARIANT 1038 1038  
 CC FT VARIANT 1198 1199  
 CC FT VARIANT 1257 1257  
 CC FT VARIANT 1392 1392  
 CC FT VARIANT 1520 1521  
 CC FT VARIANT 1583 1583  
 CC FT VARIANT 1613 1613  
 CC FT VARIANT 1686 1686  
 CC FT VARIANT 1799 1799  
 CC FT VARIANT 1881 1881  
 CC FT VARIANT 1894 1894  
 CC FT VARIANT 2141 2141  
 CC FT VARIANT 2392 2392  
 CC FT VARIANT 2605 2605  
 CC FT VARIANT 2648 2648  
 CC FT VARIANT 2717 2717  
 CC FT VARIANT 2729 2729  
 CC FT VARIANT 2814 2814  
 CC FT VARIANT 2827 2827  
 CC FT VARIANT 2907 2907  
 CC FT VARIANT 2929 2929  
 CC FT VARIANT 3058 3058  
 CC FT VARIANT 3071 3071  
 CC FT VARIANT 3081 3081  
 CC FT VARIANT 3089 3089  
 CC FT VARIANT 3105 3109  
 CC FT VARIANT 3220 3220  
 CC FT VARIANT 3238 3238  
 CC FT VARIANT 3243 3243  
 CC FT VARIANT 3245 3245  
 CC SO SEQUENCE 3329 AA; 370658 MW; AAB269APBF05A501 CRC64;

Query Match 3.7%; Score 357.5; DB 1; Length 3329;  
 Best Local Similarity 19.2%; Pred. No. 2.2e-07;



AC 000798;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE Reticulocyte binding protein 1 precursor.  
 GN RBPI.  
 OS Plasmodium vivax (strain Belem).  
 OC Eukaryote; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxID=31273;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92315338; PubMed=1617731;  
 RA Galinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.;  
 RT "A reticulocyte-binding protein complex of Plasmodium vivax  
 merozoites";  
 CC Cell 69:1213-1226(1992).  
 CC -1- FUNCTION: INVOLVED IN RETICULOCYTE ADHESION. SPECIFICALLY BINDS TO  
 CC HUMAN RETICULOCYTE CELLS.  
 CC -1- SUBUNIT: HOMODIMER (POTENTIAL).  
 CC -1- SUBCELLULAR LOCATION: Membrane-bound.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL: M88097; AAA29743.1;  
 KW Malaria; Receptor; Signal; Transmembrane.  
 FT SIGNAL 1 17 POTENTIAL.  
 FT CHAIN 18 2869 RETICULOCYTE BINDING PROTEIN 1.  
 FT DOMAIN 18 2807 EXTRACELLULAR.  
 FT TRANSMEM 2808 2826 POTENTIAL.  
 FT DOMAIN 2827 2869 CYTOPLASMIC.  
 FT SITE 1030 1032 CELL ATTACHMENT SITE (POTENTIAL).  
 FT SITE 2599 2601 CELL ATTACHMENT SITE (POTENTIAL).  
 FT SITE 2599 2601 CELL ATTACHMENT SITE (POTENTIAL).  
 SQ SEQUENCE 2869 AA: 330213 MW: 89DBE442205BECFF CRC64;  
 Query Match 3.58; Score 340; DB 1; Length 2869;  
 Best Local Similarity 17.78; Pred. No. 9.9e-07;  
 Matches 384; Conservative 340; Mismatches 742; Indels 704; Gaps 92;  
 QY 1 MDLSARVEE-VONVINAOKILECPICE-----LIKEPSTCKDHT 42  
 DB 626 IDSLSOTVEKFEYKELIDSKREKIELKEPEKSVTEINRLDODGSADLHEOIKELIDK 685  
 QY 43 FCKFCMLKLLNOKRGSPGCPKNDITKRSLOSTRFSSQVLELLKTIICAFOLDGLEXA 102  
 DB 686 AKVHYLKEKLLSKS-----SVYFTEW-NELNMT-----ASYD 719  
 QY 103 NSNFAKKNENPEHAKDEVSIISGMYNNRAKRLLOSEPNSLOETSLISVOLSMGLYV 162  
 DB 720 NMEGFSAKK---EKADNDINALYNSVYR-----EDINALIEVEKEVLENKESLT 766  
 QY 163 RLRTKRIOTPOKTSYVIELGSDSEDTYNNKATYCSVGDQELLQI-----TQGT 213  
 DB 767 EMKDKDE-----MEKKLODAKETFAKLNF--VSDDKLTDVYTKKSAEYTNAGTK 814  
 QY 214 DEIS---LDSAKKACEFSETDVTNTHQPS---NNDLNTTK---RAERHPRKYOG 263  
 DB 815 KETAOKOFENFVHKMEFSDAFSTKFEALONSMQYNOGDALTEKHKONSEKEEYFKN 874  
 QY 264 SSY---SNLHVPCGTNTHASL-----283  
 DB 875 ESVEEDLSRETEEOEYTHKKNFSRRKCEISAETNNMEVINKIESQJNYGVIEKYS 934  
 QY 284 ---OHENSLLTK-----DRMNYEAKFCNKS-----KOPGLARSGQHR 320  
 DB 935 LIDQDQEVSTARKLKEKIVSDSLRDKIDQETPFKREKTSAVENTVSTIOSLSAIDSLK 994  
 QY 321 WAGSKETCNDRTTPSTEEKVLDNADPLCERKEKNOKLP-----CSENPRDTEDEVPI 375

DB 995 LNCISNNCKKYNT-----DIDLRSKIKITLREBQKEMPRGDKGEN---TALLLKSL 1046  
 QY 376 NSSIQVNEFMSDELGSDSDHDE-----SENAKV-----ADVLDVNEVEDEYS 424  
 DB 1047 RDMRGKINE--KLNDRLNSLDTKREDLTKFYSESKSHLSKDOKGPOPLNRIDEMED 1104  
 QY 425 SSEKIDLL-----ASDPHEALICKSERVSKVEENIEDKIGKGYRKASJPNLSHY 478  
 DB 1105 IKRDVDELNVNVOISENKVYTLFKNNNSVYIEMHSHINTVANGITSNNKEILSKVEYE 1164  
 QY 479 ENLITAFYEPDIIQERPLTNKLRKRRTSGIARPEFIKADLAVOQTPEMINTGNO 538  
 DB 1165 DKL-----NLVQNEDEYKVKNPEN-----EKOLEAIRGSMKLEVIYNNHVE 1208  
 QY 539 TEONGVAMNTNG-----HE-----NKTGD-----STONEKNPPIESLE--K 576  
 DB 1209 MTQLESTANTLKKNAGKENEHDELELNKTRGQMDIYEKLLIAELEKGVNELKDN 1268  
 QY 577 ESAFKTKAEPISSISNMELELNHNSKAPK-----KRLRR 613  
 DB 1269 EKANKVEPEPERNIIQHVLEITVEKDKAGKAYVEENSLKTEKLIQETSDSOHELYT 1328  
 QY 614 KSTRIH---ALEVVSRLSPNCTELOIDSCSSSEETK-----KKRYNQPVRR 662  
 DB 1329 TSTIKHLENAGKGYEDYIKRN-----EEDSIQIREKAKSLDTLDEMKEVQV--- 1375  
 QY 663 SRLQLMECKEPTGAKKSNKPN-----QSKRHDSPTPELKLTA 705  
 DB 1376 --NMNLOSAIQGNAGISK--ELNELKQIYELLISTVYSTILEVKNSSSVREFSOLAN- 1430  
 QY 706 PGSGFTCSNTSELKEFVNPSLPREEKEKLETVKYVSNNAEDPDMLSGERVLOTGRSEY 765  
 DB 1431 -GEFTTAEG-----EKKNASARLAELAKLEQIVKRLDYSDIDDKKAKIIEGIRRELLK 1484  
 QY 766 SSSISLVPGTGYTOESISILEVSTLGAKTEPNKCVSOCAAFENPKGLHGGSKDNRD 825  
 DB 1485 ESALTWESESEKPKQCCSSHMEKAGKKIEYLK-----NN 1521  
 QY 826 TEGFKYPLGHEVHNSHETSIEME-----SELDAQVLONTFKYKSKOSFALNSPQNAEE 880  
 DB 1522 GDDGK-----ANITDSQMEVGNVYSKAAHHTVEAOVDKTKARC-----E 1563  
 QY 881 ECATFSAHSGSLKKOS--PRVTFCEQKEENQCKNSNIRKPYGVITAGFVVGQDKP 938  
 DB 1564 SIYAIVTKMDNLFNENSLAKEVKKCKEKNDDEAKYIAKLKP-----1604  
 QY 939 VDNKCSIKGSRFLCSSQFRGNETGLITPNKHGLLQNPYRIPLPIKSFVTKCKKNL 998  
 DB 1605 -----YDGRIKARVSENERKI-----SELKEKAKVE- 1630  
 QY 999 LLENFEHMSMPREKNGENIPSTVSTISRNNIRENVFKAEASSNINEGSSINEGSSI 1058  
 DB 1631 -----KRESSQJNDVSTKSLQIDNCRQOL--DSVLSNIGRAVQ-----NAL 1670  
 QY 1059 NEIGSSDENIQ-----AELGRNRG--PKLNAMLRGLVQPEVYKOSLPGSCNCKHPEIKOE 1112  
 DB 1671 QYFDSADKSMKSVLPISSELGAEKSLDKVA-----AKKS 1704  
 QY 1113 YEEVQVNTDPSPLYISDNLEQPMGS--SHASOVCETPDDL-----DQG---EIKED 1162  
 DB 1705 YERNLETVONEMSGRI---NVEE--GSLTIDIDKIIDENDLLKMKQYEEGLQIKEN 1758  
 QY 1163 TSFAEND-----IKESSAVRSKVQK-----GELSSPSP 1192  
 DB 1759 ADKRKSNFELVGSSEINALDPDSTSIPIKULKLEYDGTGLKNYKVMNEIHEFTKSYN- 1817  
 QY 1193 FTHTHL-----AGQYRGAKEKLESSPENLSDEE-----LPQO 1227  
 DB 1818 LIETHLSNATDVSVFTEKQSLRELAEK---EEHRLRREBEAFLINDIKKVESIKLK 1874  
 QY 1228 HLLFGKVNNIPOSOTRHSTVATECLSKNTEENLLSLKNSLDCS---NOVILAKASOE 1283



Db 1875 EMKRVSAEYEGMKRDHTSVSQLVODMKTYDELKTLNDISECSVLLNNVSVKVKES 1934  
 QY 1284 HLSEETKCSASFSSQCELEDTANTNTNODPFLIGSSKOMHOSQ-----GV 1333  
 Db 1935 KHADRRRANSMYEMVT-----LANYPLSDPAKISSGMEFNAEKNKTKDLELEIFSV 1989  
 QY 1334 GLSDKEVLSDDEERGTGLEENNQEBOSMDSNIGEAASGCESETSVSEDSGLSOSDILT 1393  
 Db 1990 ISNSNELLKTKIEODSNVYQKRESEQLAKDATDIYVTKLKNFERNEKLEAKNKEEYVS 2049  
 QY 1394 TQQRDITM-----HNLKLOEMALEAVLEPHGOSPS----- 1426  
 Db 2050 EKVYEAALRQSQVEGIRCHFENFHRLLDNTLEELNKKVVTYTRKKSRESGLQEMENE 2109  
 QY 1427 -NSYPS-----IISDSALEDRNPPOST-----SEKAVLSQKSE-----Y 1463  
 Db 2110 MNTYSNITTOLEGIYVAGSEKEDIKLEKRSNEMERNISEKISTSDKTYENKSTIDELY 2169  
 QY 1464 PISONPEGLSADKFEVSDSSTS-----NKEPVERSSPSKC-----PSLDDRWMHS 1512  
 Db 2170 KLGKNCQAHWISLISYTANMKTSKLMINK-----EKRETEKCVDYIKDSSSTDGYYET 2225  
 QY 1513 CSGSLQNR-NYPSOELLKVVY-----EEQLLESGPHDLTETSLPRQD-----L 1558  
 Db 2226 LKGYGSKLTFSSASAEIVQNDVYSVNFARKEKSL--NAIDIKRELTLFHQNDISIV 2283  
 QY 1559 EGTPLYEGSILFSDPDESDPESDRAPESAR-VGNIPSTSAKLPOLKVAESAQSPA 1617  
 Db 2284 EGG--VQNMALY-----DKLNEKRMDELRYRISTKLOME----- 2320  
 QY 1618 HTTDTAG-----YNAMESVSREKPELTASTERNKRM-----SNVSGL--TPEFMV 1665  
 Db 2321 HSTDFKPMIELHKMNETNNKSLKEKELKSVNDHMSMAEMIKNGLKTYTPESVQNI 2380  
 QY 1666 -----YKPAKHH-----ITLTN-----LTTEETHYV 1688  
 Db 2381 NNIVSYIEAEVKTLEIDRDYGDNYIVEEHKQESILDRTNALMDIEIKENNYNL 2440  
 QY 1689 MKTDAEFYCERLTAKFLGIAGKVV--VSFYMTQSIKERKLNHDEFGVDVNGRN 1745  
 Db 2441 MEVMTETI-HRVNDIEKIT-NKLVQAKTEYQIOLINIKQNDMLOINIFLKVSII--EX 2496  
 QY 1746 HQGPKRARES 1755  
 Db 2497 FENYKKKES 2506  
 RESULT 6  
 YD86\_SCHPO STANDARD: PRT: 1957 AA.  
 ID YD86\_SCHPO  
 AC Q10411;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Hypothetical protein C1F3.06c in chromosome I.  
 GN SPAC1F3.06c.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 NC NCBL\_TaxID=4896;  
 OX NCBL\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RX MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown S., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell J., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,

RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares K., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volckaert G., Aert R., Roben J., Grymoprez B.,  
 RA Weljens I., Vanstreels E., Rieger M., Scheifer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Jelaune V., Motier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hust S.M.,  
 RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerutti L., Lowe T., McCombie M.R., Paulsen I., Potashkin J.,  
 RA Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.;  
 RT \*The genome sequence of Schizosaccharomyces pombe\*;  
 RL Nature 415:871-880(2002).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 DR EMBL: 270690; CAA94624.1; -  
 KW Hypothetical protein. 1957 AA; 222785 MW; 3F480CA06171D9DA CRC64;  
 SQ SEQUENCE  
 Query Match 3.5%; Score 339; DB 1; Length 1957;  
 Best local similarity 18.5%; Pred. No. 6.7e-07;  
 Matches 354; Conservative 325; Mismatches 690; Indels 548; Gaps 79;  
 QY 3 LSALRVEEYQNVINA-----MOKITEPT-----CLELKEPVSTYCOHIFCFMUKL 51  
 Db 266 VSTLR--QTENSLRKECKTLQEKLEKCAINEDSKLEELKHVANYASDAIVHDKLIED 323  
 QY 52 LNOK-----KGPSQCLCKNDITKR-----SLOES--TRFSQLEELIKITICAROLD 96  
 Db 324 LSTRISEDNLSEEDTISTNEKLEKLLRNTISLSQDSKRTSNSQLEENEL----- 376  
 QY 97 TGLEAYANYPNAKKNNSPEH--LKDEYSIIQSMGYRRNARRLQSEPNLSQETSLAV 154  
 Db 377 -----KSNRTIHSQLTDAESKLSFQEQEKSLKSGSIDEYQNNLSKDKVVK 423  
 QY 155 QLSN-----LGTVRTLKQRIQOKTSVYELQSD-----SSEDTYVKAT 195  
 Db 424 QVSSQLEEARSSLAHATKLAELINSEDFQKKIKDEKLEQDRLACINSSNLEKESA 483  
 QY 196 YCSVGDOELLQTPGRTDEISLDSAKKACAFSEDTQVNTNHHQPSNNDLNTEKRAAE 255  
 Db 484 LIDKQDQELNNIREO--IKQKKVSESTOSSIQSLQDRILN-----EKR--- 525  
 QY 256 RHPEKYQGSVSNLHVEPCGTTHAASSLOHENSILLTKDRMVEKAFCNKSQPGLAR 315  
 Db 526 KH-EVYE--SQNLELKE--LQTEISNSEHLSQL----- 555  
 QY 316 SOHNRMAASKETCNDKRTPTTEKKVDLNLADPLCE--KRWKKQKLPCESENPROTEEDVPI 373  
 Db 556 ---STLAKEBAVAATNNELSEKNSLOT--LCNAPQEKLAQSMOLKENQON--FS 605  
 QY 374 TLNNSIQKVNEMFQRSDELGSDSHQSESPNKAQVADVLDVLN-----EVEYSGSSSE 427  
 Db 606 SLDTSEFKLNE--SHQEL--ENNHTITKQLKDTSSKLOQLOLERANFEQKESTLSDE 659  
 QY 428 KIDILA-----SDPEALICKSEVHSSKVSNIEDKIFGTYKK-----KASLPNL 474  
 Db 660 NNDLRTKLKLEESKSLIKKQEDV--DSLEKNIO--TLNEDKLKSEALRFLKLEKKNL 715  
 QY 475 SHVTENLITGAEVTEPQITQERPLTNKLLKRRKRPISGLHPEDFIKADLAVQKTP--EMI 532

Db 716 REVIDNL---KQHETLEAQRNDLHSSLDANK-TNALISELTKSEDEVKRLTANVEL 771  
 QY 533 NOGTNOTEON-----GQVNNITNSGHENKTKGDSIONENKPNPIESLEKESAFKAEPI 587  
 Db 772 TQDSKAMKOSFTSLVANSYOSISLHLYHELDHNMOSNN-----TLESESKLKTCENL 827  
 QY 588 SSS---ISNMEELINHNSKAPKKNRLRRKSSTRIHALELVSNLS-----632  
 Db 828 TQONMTLLINVOQLMKHYNQESKSELKEVNGKLSLDLKNLSSLNVAISDNDQILITOL 887  
 QY 633 ---PNCSTLOIDSCSSSEIKKKKKNQMPVRHSRLQ-----MEGEPATGAKSKN 683  
 Db 888 AELSKNYSLEDSQAOLNSGLASLEAKO-LHTEHEELHRLDKLTGKLKIESSSSL 946  
 QY 684 PNEOTSKRHSDTFPELKLTLNAGSEFTKSNSELEKFNPSLPREKEE---KLETVK 739  
 Db 947 GKLTJROEE---ISMLKEENMSQSAITSVKSLDETLSKSLKLEADIEHLKNKXSEVE 1003  
 QY 740 VSNNAEDPKDMLSGERVLTQTERSVSSISLVPGTDTQESISLLEVTGKAKTEPN 799  
 Db 1004 VERNL-----LLASNERLMD-----1018  
 QY 800 KCVSOCAAFENPKGLIHGSKDNNDTEGKYPPLGHEVNHRSRTSIEMESELDAQYLO 859  
 Db 1019 -----DLKNNGENIA-----SLQTELEKRAENDD--LOS 1046  
 QY 860 TFKV--SKROSFALFENPGNAEPECATFSAHSGSLKQSPKVPFEQKEENCK--NNS 915  
 Db 1047 KLSVSVSEYENLLISSQTN-----KSLBKTQNKLYTEKNQKLLDK 1090  
 QY 916 NIKPVQVNTAGFPVVGQDKRVDNAKCSIKGSRFCLSSQFRNGETGLTPNKHGLQ 975  
 Db 1091 DQHNVELLETSKYKGLGEENQIKDELLALRKSK-----KQHDLOA 1133  
 QY 976 NPIRIPPLPPISEFVTKCKKNLLEENFEHNSMPRENGENIPISTVSTISRNITRENV 1035  
 Db 1134 N-----FV-----DLEKESDALE-QLTENKELIVLEDSNNSNEAL 1170  
 QY 1036 FKEAS--SSNINEVGSTNEVGSSINEIGSSDENIDAEGRNRPKLMRLGLVQPPV 1093  
 Db 1171 VERISDLARLSDMKKSLSDSNVIVISRLDVRVDEL-----DT 1211  
 QY 1094 YKQSLPGSNCKHPEIKQYEEVQTV--NTDFSPYLLISDNLQOPMGSSHASQVCEPT 1150  
 Db 1212 LKKRDSLSTQYSEV-CQDRDLDSLCKCESEFNKVAVS-----LRELCTKSE 1259  
 QY 1151 DD-----LDDGEIKEDTSFA-----ENDIKESAV-FSKVQKGLSRSPPT 1194  
 Db 1260 IDVPVSEILDNVEFVAGNFSELSRLTVLSLENYLDAFQVNEFKKMLDNRLLTTDAEPT 1319  
 QY 1195 -----HTHLAGYRRG--AKKLESSEENLSEDELEPCFQHLFGVNNIPSO 1242  
 Db 1320 KYVADLEKLOHEHMDLQROGLEKALKSEKNFLKREKM-----TENHLSLE 1369  
 QY 1243 RSHVATEC--LSKNTENLILS--LKNSLDCSNQVILAKASQEHLSSEETKCSAELS 1297  
 Db 1370 GKRETKETIAELSSRLDQNLATNKLKNQD-----HLMOETRLKRDVAK 1414  
 QY 1298 SOCS---ELEDTLANTQDPLIGSSKOMRHOSESQGVLSDELVSDEBERGICLEN 1354  
 Db 1415 EKESLLISLEESLSNOROKESSLLDAKNELEH-----MLDTSRNSLSMK 1461  
 QY 1355 NOE-EQSMDSNLGEAAGSESETSVSEDCSGLSSOSDILTQOQRTDMONMLKLOEME 1413  
 Db 1462 IBSINSLLDKSKSELASAVEKIGALOK-----LHSELSLMENTKSQLQAKERIOVDE 1517  
 QY 1414 LEAVLEQHSQPSNSY-----PSLIISD-SSALEDLRN--POSTSEKAVILTSOKSSEY 1463  
 Db 1518 IOE-LDHEITASKNNYEGLKNDKSIIRLUSENIEQNLMLAEKSAVAVRLSTPEKSEIL 1576  
 QY 1464 PISQNEGSLADKFEVSADSSTSKNKEPVERSSPKCSLDDRWYMHSCSSGLQNRNP 1523  
 Db 1577 QFNSRLADLEHYHKSQVSESELGRSKLK-----LASTTELEQLAENE 1616

QY 1524 SOEELIKVVDVEEOQLESGBDLETETSYLPRODLEGTPLYESGIS-----1569  
 Db 1617 RLSTLRMDLQO-----VKDLSNIKDLSLSDLTLSLEBSVASLOKECKIKSNYVE 1670  
 QY 1570 ----LFSDDPESDPESDRAPESA-----RVGNIPSTALKVPQKLVASASPAHAHT 1619  
 Db 1671 SLQDVLTYSQVARNAELEDEVSRSVDKIRRRDDCEHLSCGLKHLHSQLQLEQHEFFRAEQ 1730  
 QY 1620 TDTAGYNAMESYRREKPELTASTERVNRKMSVNSGLPPEEMLYKFF-----1668  
 Db 1731 QRRTOGLFLETKYKKO-----EKLLKLNLRQEQLLPRSSILVYESYIRDIKEIIV 1782  
 QY 1669 --ARKHHTLTLNLTIT-----ETHVVKMTAEFCERTLKLYFLGAGKVVASY 1716  
 Db 1783 LQERLNGIELSQLPRGICYGYFFKTRVEMEVLDSKQOYAKIQFL--AGAETIVAF 1837

RESULT 7  
 MLPL YEAST  
 ID MLPL YEAST STANDARD: PRT: 1875 AA.  
 AC 002455;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Myosin-like protein MLPL.  
 GN MLPL OR YKR095W OR YKR415.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=5288C;  
 RX MEDLINE=93247549; PubMed=8483450;  
 RA Koelling R., Nguyen T., Chen E.Y., Botstein D.;  
 RT "A new yeast gene with a myosin-like heptad repeat structure.";  
 RL Mol. Gen. Genet. 237:359-369(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94205265; PubMed=8154186;  
 RA Bou G., Esteban P.F., Baladron V., Gonzalez G.A., Cantalejo J.G.,  
 RA Remacha M., Jimenez A., del Rey F., Ballesta J.P.G., Revuelta J.L.;  
 RT "The complete sequence of a 15,820 bp segment of Saccharomyces  
 cerevisiae chromosome XI contains the UBI2 and MPL1 genes and three  
 RT new open reading frames.";  
 RL Yeast 9:1349-1354(1993).  
 CC  
 CC -1- FUNCTION: MYOSIN-LIKE PROTEIN THAT IS PROBABLY INVOLVED IN DNA  
 CC REPAIR.  
 CC  
 CC -1- SIMILARITY: SOME, TO THE TPR ONCOGENE.  
 CC -1- CAUTION: REF. 2 MISQUOTES THE GENE NAME AS "MPL1".  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: L01992; AAA34783.1;  
 CC EMBL: X73541; CAA51948.1;  
 CC EMBL: Z28320; CAA82174.1;  
 CC PIR: S38173; S38173.  
 CC DR SGD: S0001803; MPL1.  
 CC Coiled coil; DNA repair.  
 CC KW Coiled coil; DNA repair.  
 CC FT DOMAIN 69 487  
 CC FT DOMAIN 531 1678  
 CC FT DOMAIN 1834 1866  
 CC FT CONFLICT 301 301  
 CC FT SEQUENCE 1875 AA: 218455 MW: 68340D34C906867 CRC64;  
 Query Match 3.4%; Score 327; DB 1; Length 1875;

Best Local Similarity 18.7%; Pred. No. 2e-06;  
Matches 300; Conservative 269; Mismatches 603; Indels 432; Gaps 70;

4 SALRVEEYQVNTINAM-----OKILEPCHELKEPVSTKCHDHFCKFCMLKLNK 55  
438 AALLLEHSTNEKNAKVELNAKNOKEVENDLQITTKORLDICROLYLLITNSVNS 497  
56 KGP-----SQCPLCKNDITKRSLOESTRPSOLVE-----ELKII--CA 92  
498 KGPLKEEIOFIQNIQEDDSTITESDQKVTERLVEFKNIIQOEKNAELLKVRNLA 557  
93 FOLDGLLEYANSYNAKKNENSPHILKDVSTIIS--WGYRNRAKLLQSEPNLSLOET 150  
558 DKLES--KERKSKOSLOKIESEFTVNAKEAIIITLSEKMDLESRIE--LOKELEELKTSVP 615  
151 SLVQSLNIGVRLTKORIOPO-----KTSV 178  
616 NEDASYSVNTTKQLETREKDESQVODIOTRISQITRESTENMSLNKEIOLDYKSDI 675  
179 YIELGSDS-----EDTVNKA 194  
676 SIKIGKEKSRILAEERFKLSNTLDLTKAENDQLRKRDYLONTLKQDSKTHETLNEY 735  
195 TYC-----SVGOELLQITPO-----GTRDEISLSQAKACEESETDVYTNHHP 241  
736 VSCSKSLIVETELLNKEEORLKVHLEKNLQOEINLKSPKDSLRIM---VTOLQITOK 792  
242 SNNIDNTEKRAEHPERKYGSSVSNLHVEPGNTNTHASSLOHENSLLTKDRNVE- 300  
793 EREDLEETRKSCQCKIDELF--DALSELKKEITSQDHIKOLEEDNNS-----NIW 843  
301 ---KAEPCKN-----SKQPLASQHNRAWAGSEKTCNDRTPTSEKVD--LNAD 345  
844 YONKIEALKKYEESVITSVSKQTDIEKQYKVKLEKEIEEDKIRLHTYNNMDEITND 903  
346 PLCEKKEKNOKLPCE--NPRDEYDPMITLNSIQVNMFRSDLLG----- 394  
904 SL--RKELEKSKINLIDAYISOIKKYDL--YETTSOLOOTN---SKLDSFKOFTNQIKN 957  
395 -SDSHDESESNKAVADVLNLEVD-----EYSGSESEKIDLASDPEALICKSE 445  
958 LTDEKTSLEDKISLKEQMFNLNNELDLQKGMKEKADFKRKISILONNKEVAVNSE 1017  
446 -----RVHSKSVESIEDKIFEGYRKKASL--PNLSHTENLIIGAFYTE 489  
1018 YESKLSKIQNDLDOITIANTAQNNYEOEL-----OKHADVSKITSELREQHTYGGQYK 1072  
490 POIIQERLTKKRRRPTGSLHPEDIRKADLAVQCTPEINQCTNOTBQNGVMTT 549  
1073 TLNLSROLENALENKSWSS--QKESLLEDLDSNRIEDLSSQNKLLYDQI--QIYTA 1130  
550 NSGHEKTKGDSION-----EKNPPIESLEKESAFKTKAEPISISSIMLEL-----N 599  
1131 DKEVNNSTNGPGLNILLITLRERDIDT--KVTAEDAKMLRKISIMDELQDARTK 1188  
600 IHNSKAPKRN---LRRKSTRIHILELVVSRNLSPNCTEQIDSCSSE---EIKK 652  
1189 LDNSRVEKENSIIQCHDDIMEKLNQNLRLRESNITRN--ELENNNNKKKEIQSELDK 1246  
653 KKYQMPVRHSRNQIMGKEPATGAKSKNPNQETSRRHSDTPEPLKLTNAPSFTKC 712  
1247 LKQVAVP-----ESELTAIKYSQEKEDQ-----ELKLAK----- 1276  
713 SNTSELKEFVNPSPREKEKELET---VKVSNNADEPKDMLSGERY----- 757  
1277 ---BEVHRMKKRQDIILEKHQSSDOVEKLESEIENLKELENNKRGAEAEKFNLR 1333  
758 LQTRRSVSESSISLVPGTDYGTQESISLLEVSTLGAKTAPNKCVSQCA---FENPKG 813  
1334 ROADERKTSKLS-----QDSLT--EQVNSLTDANKVENSISEANARIEELQNK- 1382  
814 LIHCSMDNRNDTGFKYPLGHEVNHRSRETSIEMSESLDAQYLQ-----TFK--VS 864

1383 -----VAQNNOLEAIIRKLOEDAERKASRELQAKLEESTTSYESTINGINEIITLKEEIE 1437  
865 KROSF--ALFNSPNAEEPCATFSAHSGSLKK--OSPVTPECEQKEPNQKNSNKPV 920  
1438 KQROIQOQLOATISANDQNSLNI--SMKSFEDKIKF---IKEKTOEVNKEIIEAQ 1491  
921 QTVNITAGFPVVGOKDKPYDNKCSIKGSRFCLSSQPRGNETGLTPNKHGLQNPYRI 980  
1492 ERLNDQNSINIMEETIKKMESEHDEY-----SQKIREAEAA-----LKKRI 1534  
981 PPLFPKTSVTKCKKNLLENEFESHSPERKNGENIPSTVSTISNNIR----- 1032  
1535 PTEERIKNIIEER--KKEELEKEFEED--KVEERIKSMQOSEIIVVLKOLEAKVOEKOK 1589  
1033 --ENVFKEASSNINEVGSSTN-----EVGSSINGISSDENIQALGRNPK 1079  
1590 ELENEYNKKLOEELDVPHSSITSDEDEDKLRALIESKRLKEEPNNELOAIKKSFDG-K 1648  
1080 LNAMLRLGVLOPEVYKOSLPGSNCKHPKIKQOEYEVQVNTDPSYLLSDNLEOP-MG 1138  
1649 QQAMMKTTLLERKLAKMESQJSE-----TQASAPSPKSVN-----NVQNPILG 1692  
1139 SSHASQVSETPDDLLDGEIKEDTSEFANDIKESSAVFSKSVQKGLSRSPEFTHTL 1198  
1693 IPRKIEENSNSPFPNLLSGE-----KLKLNKSSSGGCFNFT-----SPSPNKH-- 1737  
1199 AQGYRGAKKLESSEENISSEDELPQFQHLHFGVNNIPQS---TRHSTVATECLSKN 1255  
1738 ---LQNDNDKRESLANKNDPPLHLEPSF-----NTPASRLISSSTLSD---T 1781  
1256 TEENLSEKNSLNDQSNQVITAKASQ-----HMLSETKCSALFFSQSCEL- 1303  
1782 NDEEFTSNPQAKDSSNBNVQSEEDTEKKKEGEPVKRGEALIEQTKSNKRP-DEVGELK 1840  
1304 ---EDLTANTNQPDPFLGSSQKMRHQSQSGVGLSKELVSD 1344  
1841 NDEDJTENIN-----ESKRIKTEDEE---KETDKVND 1872

RESULT 8  
BRC2\_HUMAN  
ID BRC2\_HUMAN STANDARD. PRT. 3418 AA.  
AC P51587; Q1879; O00183; O15008;  
DT 01-OCT-1996 (Rel. 34, Created)  
DI 16-OCT-2001 (Rel. 40, Last sequence update)  
DR 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Breast cancer type 2 susceptibility protein.  
GN BRC2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
OX NCBI\_TaxId=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96112016; PubMed=8524414;  
RA Wooster R., Bignelli G., Lancaster J., Swift S., Seal S., Mangion J.,  
RA Collins N., Gregory S., Gumbs C., Micklem G., Barfoot R., Hamoudi R.,  
RA Patel S., Rice C., Biggs P., Hashim Y., Smith A., Connor F.,  
RA Arson A., Gudmundsson J., Filene D., Kelsell D., Ford D., Tonin P.,  
RA Bishop D.T., Spurr N.K., Fonder B.A.J., Beles R., Peto J., Devilee P.,  
RA Coriell C., Lynch H., Narod S., Lenoir G., Egilsson V.,  
RA Barkodt R.B., Easton D.F., Bentley D.R., Futreal P.A.,  
RA Ashworth A., Stratton M.R.;  
RT "Identification of the breast cancer susceptibility gene BRC2";  
RL Nature 378:789-792(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96172838; PubMed=8589730;  
RA Tavtigian S.V., Simard J., Rommens J., Couch F., Shattuck-Eidens D.,  
RA Neuhansen S., Merafyer S., Thorlacius S., Offit K., Stoppa-Lyonnet D.,  
RA Belanger C., Bell R., Berry S., Bogden R., Chen Q., Davis T.,  
RA Dumont M., Frye C., Hattier T., Jamulapati S., Janacki T., Jiang P.,  
RA Kerner R., Leblanc J.-F., Mitchell J.T., McArthur-Morrison J.,

RA Nguyen K., Peng Y., Samson C., Schroeder M., Snyder S.C., Steele L.,  
 RA Stringfellow M., Strop C., Swedlund B., Swensen J., Teng D.,  
 RA Thomas A., Tran T., Tran T., Tranchant M., Weaver-Feldhaus J.,  
 RA Wong A.K.C., Shizuya H., Eyford J.E., Cannon-Albright L., Labrie F.,  
 RA Skolnick M.H., Weber B., Kamb A., Goldar D.E.,  
 RT "The complete BRCA2 gene and mutations in chromosome 13q-linked  
 RT kindreds.";  
 RL Nat. Genet. 12:333-337(1996).  
 RN [13]  
 RP SEQUENCE FROM N.A.  
 RA Hunt S., Murray A., Williamson H.,  
 RL submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.  
 RN [14]  
 RP VARIANT OC HIS-2787, AND VARIANTS ASN-372; MET-1915 AND VAL-2466.  
 RX MEDLINE=96275740; PubMed=665505;  
 RA Takahashi H., Chiu H.-C., Bandera C.A., Behrhardt K., Liu P.C.,  
 RA Couch F.J., Weber B.L., Livolsi V.A., Furusato M., Rebani B.A.,  
 RA Cardonick A., Benjamin I., Morgan M.A., King S.A., Mikuta J.J.,  
 RA Rubin S.C., Boyd J.,  
 RT "Mutations of the BRCA2 gene in ovarian carcinomas.";  
 RL Cancer Res. 56:2738-2741(1996).  
 RN [15]  
 RP VARIANTS ASN-372; ASP-991; SER-1147; MET-1915 AND CYS-2034.  
 RX MEDLINE=96241588; PubMed=6673091;  
 RA Couch F.J., Farid L.M., Deshano M.T., Tavtigian S.V., Calzone K.,  
 RA Campeau L., Peng Y., Bogden B., Chen O., Neuhausen S.,  
 RA Shattuck-Eidens D., Godwin A.K., Daly M., Radford D.M., Sedlacek S.,  
 RA Rommens J., Simard J., Garber J., Merajver S., Weber B.L.,  
 RT "BRCA2 germline mutations in male breast cancer cases and breast  
 RT cancer families.";  
 RL Nat. Genet. 13:123-125(1996).  
 RN [16]  
 RP VARIANT GLU-3095.  
 RX MEDLINE=96225456; PubMed=8640235;  
 RA Lancaster J.M., Wooster R., Mangion J., Phelan C.M., Cochran C.,  
 RA Gumbs C., Seal S., Barfoot R., Collins N., Bignell G., Patel S.,  
 RA Hamoudi R., Lattson C., Wiseman R.W., Berchuck A., Iglehart J.D.,  
 RA Marks J.R., Ashworth A., Stratton M.R., Futreal P.A.,  
 RT "BRCA2 mutations in primary breast and ovarian cancers.";  
 RL Nat. Genet. 13:238-240(1996).  
 RN [17]  
 RP VARIANTS  
 RX MEDLINE=96225457; PubMed=8640236;  
 RA Teng D.H.-F., Bogden R., Mitchell J., Baumgard M., Bell R., Berry S.,  
 RA Davis T., Ha P.C., Kehrner R., Jammulapati S., Chen O., Offit K.,  
 RA Skolnick M.H., Tavtigian S.V., Jhanwar S., Swedlund B., Wong A.K.C.,  
 RA Kamb A.,  
 RT "Low incidence of BRCA2 mutations in breast carcinoma and other  
 RT cancers.";  
 RL Nat. Genet. 13:241-244(1996).  
 RN [18]  
 RP VARIANT ASN-2415.  
 RX MEDLINE=96225458; PubMed=8640237;  
 RA Miki Y., Katagiri T., Kasumi F., Yoshimoto T., Nakamura Y.,  
 RT "Mutation analysis in the BRCA2 gene in primary breast cancers.";  
 RL Nat. Genet. 13:245-247(1996).  
 RN [19]  
 RP VARIANT BC ASP-2089, AND VARIANT VAL-3412.  
 RX MEDLINE=97294396; PubMed=9150152;  
 RA Veinanen P., Friedman L.S., Eerola H., Sarantausta L., Pylhonen S.,  
 RA Ponder B.A.J., Muhoonen T., Nevalinna H.,  
 RT "A low proportion of BRCA2 mutations in Finnish breast cancer  
 RT families.";  
 RL Am. J. Hum. Genet. 60:1050-1058(1997).  
 RN [10]  
 RP VARIANT BC AND PANCREAS CANCER TRP-554.  
 RX MEDLINE=98316775; PubMed=9654203;  
 RA Ganguly T., Dhulipala R., Godmilow L., Ganguly A.,  
 RT "High throughput fluorescence-based conformation-sensitive gel  
 RT electrophoresis (F-CGSE) identifies six unique BRCA2 mutations and an  
 RT overall low incidence of BRCA2 mutations in high-risk BRCA1-negative  
 RT breast cancer families.";  
 RL Hum. Genet. 102:549-556(1998).

RN [11]  
 RP VARIANTS BC L-32; R-53; L-81; R-201; A-211; S-222 AND T-3118.  
 RX MEDLINE=98272917; PubMed=9609997;  
 RA Katagiri T., Kasumi F., Yoshimoto M., Nomizu T., Asaishi K., Abe R.,  
 RA Tsuchiya A., Sugano M., Takai S., Tameda M., Fukutomi T., Naha K.,  
 RA Makita M., Okazaki H., Hirata K., Okazaki M., Futusuma Y.,  
 RA Morishita Y., Iino Y., Karino T., Ayabe H., Hara S., Kajiwara T.,  
 RA Hougou S., Shimizu T., Toda M., Yamazaki Y., Uchida T., Kunitomo K.,  
 RA Sonoo H., Kurebayashi J.-T., Shimotsu K., Nakamura Y., Miki Y.,  
 RT "High proportion of missense mutations of the BRCA1 and BRCA2 genes in  
 RT Japanese breast cancer families.";  
 RL J. Hum. Genet. 43:42-48(1998).  
 RN [12]  
 RP VARIANTS OC PRO-75; HIS-2502 AND HIS-3098.  
 RX MEDLINE=20455732; PubMed=10486320;  
 RA Gayther S.A., Russell P., Harrington P., Antoniou A.C., Easton D.F.,  
 RA Ponder B.A.J.,  
 RT "The contribution of germline BRCA1 and BRCA2 mutations to familial  
 RT ovarian cancer: no evidence for other ovarian cancer-susceptibility  
 RT genes.";  
 RL Am. J. Hum. Genet. 65:1021-1029(1999).  
 RN [13]  
 RP VARIANTS HIS-289; ASN-372; ASP-991 AND VAL-3412.  
 RX MEDLINE=99254821; PubMed=10323242;  
 RA Li S.-L., Tseng H.-M., Yang T.-P., Liu C.-H., Teng S.-J.,  
 RA Huang H.-W., Chen L.-M., Kao H.-W., Chen J.H., Tseng J.-N., Chen A.,  
 RA Hou M.-F., Huang T.-J., Chang H.-T., Mok K.-T., Tsai J.-H.,  
 RT "Molecular characterization of germline mutations in the BRCA1 and  
 RT BRCA2 genes from breast cancer families in Taiwan.";  
 RL Hum. Genet. 104:201-204(1999).  
 RN [14]  
 RP VARIANTS BC, AND VARIANTS.  
 RX MEDLINE=99138688; PubMed=9971877;  
 RA Wagner T.M.U., Hirtelheuer K., Shen P., Moeslinger R., Muhr D.,  
 RA Fletschmann E., Concin H., Doeller W., Haid A., Lang A.H., Mayer P.,  
 RA Petru E., Ropp E., Langbauer G., Kubista E., Scheiner O.,  
 RA Underhill P., Mountain J., Stierer M., Zielinski C., Oefner P.,  
 RT "Global sequence diversity of BRCA2: analysis of 71 breast cancer  
 RT families and 95 control individuals of worldwide populations.";  
 RL Hum. Mol. Genet. 8:413-423(1999).  
 CC -i- FUNCTION: MAY PARTICIPATE IN A PATHWAY ASSOCIATED WITH THE  
 CC ACTIVATION OF DOUBLE-STRAND BREAK REPAIR AND/OR HOMOLOGOUS  
 CC RECOMBINATION.  
 CC -i- SUBUNIT: INTERACTS WITH RAD51.  
 CC -i- TISSUE SPECIFICITY: HIGHEST LEVELS OF EXPRESSION IN BREAST AND  
 CC THYMS, WITH SLIGHTLY LOWER LEVELS IN LUNG, OVARY, AND SPLEEN.  
 CC -i- DISEASE: BREAST CANCER (BC) IS AN EXTREMELY COMMON MALIGNANCY,  
 CC AFFECTING ONE IN EIGHT WOMEN DURING THEIR LIFETIME. A POSITIVE  
 CC FAMILY HISTORY HAS BEEN IDENTIFIED AS MAJOR CONTRIBUTOR TO RISK OF  
 CC DEVELOPMENT OF THE DISEASE, AND THIS LINK IS STRIKING FOR EARLY-  
 CC ONSET BREAST CANCER. MUTATIONS IN BRCA2 ARE THOUGHT TO BE  
 CC RESPONSIBLE FOR SOME INHERITED BREAST CANCER. IT IS LINKED WITH  
 CC MALE BREAST CANCER.  
 CC -i- SIMILARITY: CONTAINS 8 BRCA2 REPEATS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: X95152; CA64484.1; JOINED.  
 DR EMBL: X95153; CA64484.1; JOINED.  
 DR EMBL: X95154; CA64484.1; JOINED.  
 DR EMBL: X95155; CA64484.1; JOINED.  
 DR EMBL: X95156; CA64484.1; JOINED.  
 DR EMBL: X95157; CA64484.1; JOINED.  
 DR EMBL: X95158; CA64484.1; JOINED.  
 DR EMBL: X95159; CA64484.1; JOINED.  
 DR EMBL: X95160; CA64484.1; JOINED.  
 DR EMBL: X95161; CA64484.1; JOINED.





CC -!- CAUTION: REF. 6 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO TWO  
 CC FRAMESHIFTS IN POSITIONS 3782 AND 3811.  
 CC -!- CAUTION: REF. 9 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO FOUR  
 CC FRAMESHIFTS IN POSITIONS 29, 1653, 1659 AND 1735.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is: in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AJ131693; CAB40713.1; -  
 DR EMBL; AB019691; BAA78718.1; -  
 DR EMBL; AJ010770; CAA09361.1; -  
 DR EMBL; AF026245; AAB6384.1; -  
 DR EMBL; AF083037; AAD22767.1; -  
 DR EMBL; AC004013; AAB96867.1; ALT\_FRAME.  
 DR EMBL; AF091711; AAD39719.1; -  
 DR EMBL; AB018346; BAA34523.1; -  
 DR EMBL; AC000066; AAC60380.1; ALT\_FRAME.  
 DR Genew; HGNC:379; AKAP9.  
 DR MIM; 604001; -  
 DR Colled coil; Alternative splicing: Polymorphism.  
 KM DOMAIN 2534 2567 PNA-RIT SUBUNIT BINDING DOMAIN.  
 FT DOMAIN 164 914 COILED COIL (POTENTIAL).  
 FT DOMAIN 944 1022 COILED COIL (POTENTIAL).  
 FT DOMAIN 1100 1185 COILED COIL (POTENTIAL).  
 FT DOMAIN 1253 1280 COILED COIL (POTENTIAL).  
 FT DOMAIN 1336 1382 COILED COIL (POTENTIAL).  
 FT DOMAIN 1434 1459 COILED COIL (POTENTIAL).  
 FT DOMAIN 1585 1659 COILED COIL (POTENTIAL).  
 FT DOMAIN 1857 2455 COILED COIL (POTENTIAL).  
 FT DOMAIN 2544 2561 COILED COIL (POTENTIAL).  
 FT DOMAIN 2603 2776 COILED COIL (POTENTIAL).  
 FT DOMAIN 3065 3092 COILED COIL (POTENTIAL).  
 FT DOMAIN 3124 3470 COILED COIL (POTENTIAL).  
 FT DOMAIN 3587 3689 COILED COIL (POTENTIAL).  
 FT DOMAIN 3726 3730 COILED COIL (POTENTIAL).  
 FT DOMAIN 203 292 GLN-RICH.  
 FT DOMAIN 321 1010 GLU-RICH.  
 FT DOMAIN 1846 2772 MISSING (IN ISOFORM 2 AND ISOFORM 3).  
 FT VARSPLC 17 28 MISSING (IN ISOFORM 4).  
 FT VARSPLC 1637 1642 MISSING (IN ISOFORM 3).  
 FT VARSPLC 1643 3911 MISSING (IN ISOFORM 4).  
 FT VARSPLC 2175 2182 MISSING (IN ISOFORM 3).  
 FT VARSPLC 2175 2183 MISSING (IN ISOFORM 3).  
 FT VARSPLC 2855 2907 SADTFKQVE -> Q (IN ISOFORM 6).  
 FT VARSPLC 2855 2907 VEFYINMCESTLC -> GSSIPELASDAVOTREICSS  
 FT VARSPLC 2895 2948 MISSING (IN ISOFORM 2, ISOFORM 3 AND ISOFORM 6).  
 FT VARSPLC 3901 3911 MISSING (IN ISOFORM 5).  
 FT VARSPLC 3901 3911 SPTOPHAGMR -> ALSLTSMOHSARPTAPLPEILSH  
 FT VARSPLC 3901 3911 SLG (IN ISOFORM 6).  
 FT VARIANT 1347 1347 K -> KO.  
 FT VARIANT 1347 1347 /FTIG-Var\_010926.  
 FT CONFLICT 76 76 E -> Q (IN REF. 3).  
 FT CONFLICT 475 475 M -> I (IN REF. 3).  
 FT CONFLICT 554 554 E -> G (IN REF. 3).  
 FT CONFLICT 638 638 R -> S (IN REF. 3).  
 FT CONFLICT 663 663 N -> S (IN REF. 3).  
 FT CONFLICT 913 913 H -> N (IN REF. 3).  
 FT CONFLICT 956 956 K -> N (IN REF. 3).  
 FT CONFLICT 980 982 OKH -> PKP (IN REF. 1 AND 2).  
 FT CONFLICT 997 997 Q -> P (IN REF. 1 AND 2).  
 FT CONFLICT 1001 1001 Q -> P (IN REF. 1 AND 2).  
 FT CONFLICT 1020 1020 N -> D (IN REF. 3).  
 FT CONFLICT 1028 1028 V -> E (IN REF. 3).  
 FT CONFLICT 1626 1626 R -> P (IN REF. 1 AND 2).  
 FT CONFLICT 1703 1703 N -> T (IN REF. 3).  
 FT CONFLICT 1707 1707 V -> G (IN REF. 3).  
 FT CONFLICT 1802 1803 MISSING (IN REF. 5).  
 FT CONFLICT 1843 1843 A -> P (IN REF. 3).  
 FT CONFLICT 1956 1956 I -> V (IN REF. 3).

FT CONFLICT 2027 2027 V -> D (IN REF. 5).  
 FT CONFLICT 2157 2158 EI -> HE (IN REF. 7).  
 FT CONFLICT 2169 2169 E -> V (IN REF. 3).  
 FT CONFLICT 2514 2169 L -> R (IN REF. 3).  
 FT CONFLICT 2514 2514 L -> R (IN REF. 3).  
 FT CONFLICT 2851 2851 I -> N (IN REF. 8).  
 FT CONFLICT 2957 2957 E -> D (IN REF. 3).  
 FT CONFLICT 2983 2983 P -> S (IN REF. 3, 7 AND 8).  
 FT CONFLICT 3087 3087 Q -> H (IN REF. 3).  
 FT CONFLICT 3218 3218 Q -> H (IN REF. 3).  
 FT CONFLICT 3307 3309 ESE -> OSO (IN REF. 3).  
 FT CONFLICT 3751 3751 P -> A (IN REF. 3).  
 FT CONFLICT 3833 3833 T -> S (IN REF. 3).  
 SQ SEQUENCE 3911 AA: 453664 MM: 3FBICBIC18B47AA CRC64;  
 Query Match 3.4%; Score 325; DB 1; Length 3911;  
 Best Local Similarity 18.0%; Pred. No. 6; 3e-06;  
 Matches 387; Conservative 338; Mismatches 756; Indels 670; Gaps 92;  
 QY 7 RVEGVNIVANQKILCEPICLEIKPEVSTKCDHIFKFCMLKLLNKKGPSQPLCKN 66  
 DB 1351 QVQELSLISSIQ-----QQLKTEQNYEAELHCKRLQAVSESTVPPSLPVDV 1401  
 QY 67 DITKRSLQSTREFSOLVEELKILICAFQDQLQLEVANYFRKKN-----SPHL 118  
 DB 1402 VITESDAQRTMYPGSCVK-----NIDGTIEFSGEG-VKEETNIVKLEKQYQYQOL 1452  
 QY 119 KDEVS-IIOSM-----GYENRA---KRLQSEPNPSIQETSLSYQSLNL 159  
 DB 1453 EEEVAVKIVISMSIAFAQOTELSRISGKENTSSQAAVQOQBEHYENKLSQDOJGF 1512  
 QY 160 GTVVRT--LTKRQIDQKTSVYIEIG-----SSSEDTYKATYCSVGQDQLQ- 206  
 DB 1513 QFFETVDVAFKKEEPLSK-----ELGEHGKEILLNSPDHIDPEKDCVLITSEMFESKD 1568  
 QY 207 ---ITPGQTRDEL---SIDSAKKAQFSEPD-----VTNHEHPNSNNDLNTERRAA 254  
 DB 1569 KTFIVRQSLHDEIVSSMDASRQMLNEBQEDMRQELVROEHOATTELLROAHNMQW 1628  
 QY 255 EHPKERYQSSVSNLHVEPCGTN---THASSLOHNSSLTLTKDRMNEYKAFCNKSOP 311  
 DB 1629 EQREDEQ-----QLOELIKRLNQLAQRSSIDNEN--LVSRENVLLLELE---ALKL 1678  
 QY 312 GLARSOHNWAGSKETCDNRTPSTEEKYVDLADPLCEKKEKNNOKLPSCSNPDTEDVP 371  
 DB 1679 SL-----AGREKLCELRNNSQTQ---NGNE--NGQVEDEQTFKEKELDRKPDVP 1725  
 QY 372 WITLNS--SIQKV-----EFMSDELGSDHDSDESNKAV- 409  
 DB 1726 PETLSERYALQAKANNRLKILLEVYKTTAAVEETIGR--HYLGILD--RSSKSSQSSASLI 1782  
 QY 410 -----ADVLDVINE-----VDEYSGSS---EKIDLLASDPHEALICKSERVHS- 449  
 DB 1783 WSESEASVSKSVHHEHFRVTDESLPSYSGSDMPRNDLNMMSKYTEBETELSOQLVRSFG 1842  
 QY 450 --KSVESNIEDKIFGKYTKRKASLNLSHV---TENLITGAFVTEPQIT---QERPLT 499  
 DB 1843 AGTEIDPENEEELMLNISRLQAAVEKLEALISETSSQLEHKAQVOTELMRESFQKQEA 1902  
 QY 500 NKLKRRRTSGLHPED-----FIKAD-----LAVOKT--PEMINOQT---NOTEQ 541  
 DB 1903 ESLKQCEELREKLHEESRAREQLAVELSKAEGVIDGYADEKTLPEEROIQETDITIDRQ 1962  
 QY 542 NQGVANNITNSGHENTKQDSIQNEKNPNPISLEKSAFKTKAEP-----SSSISM 594  
 DB 1963 --ELICASRLQELAEBOQIOEBE---LLSRQKE-AMKADAGVBOQLLOQETFKLMKE 2016  
 QY 595 ELETNINHSKADKKRRLKRSSTRIHLELVVSNLSPPNCTELQIDSCSSSEIIRKKK 654  
 DB 2017 KLEVCOQAEKV-----RDLOKOVKALEIDIVEBOVS----- 2047  
 QY 655 VNOMVVRHSRNQLMEGEPATGAKSKPK-----NEQT--SKRDSQTFP-----ELKL 702  
 DB 2048 -RFIELEQKNTLMDLRQONQALQKLEKMRKFLDEQADIDREHDFVFOEIQLOKLEOL 2106



```

QY 703 TNAPGSFTKCS--NTSELKEFPVNPSPLEPREKEEKLFTVKVSNNAEDPKDMLSGERYLQT 760
D 2107 KVP-PROPISEHOTREVEQLAN---HLKERTDKCSLLLSKE-QLOORDQERNEEKL 2161
QY 761 ERSVESSISLVPTDYGTQESISLLEVLTLGAKAKEPKNCVOCOAFAENPKGLIHGCSK 820
D 2162 EFRRELEOALIVSAD--TFQKVE--DRKHFGAVEAKPE--LSLEVOLOQAEPAIDREK 2215
QY 821 DNRDTEGFPYPLGHEVNSHRETSIEMESELDAQYQNTFVKVSKRSPFLFSPGNAE 880
D 2216 EITN-----LEEOLEQFRE--ELEKNNEEVOQLHMOLETOCKESTRLO--ELEQ 2261
QY 881 ECATFS-----AHSGLKQSPKVTPE-----CEQKE-ENQKNSNIKPVQTVN 924
D 2262 ENKLFKDMKELGLAITKESDAMSTQDOHVLFGKFAQITQKEVEYIDQNLQVYKQLOQK 2321
QY 925 ITAGFPVVGOKKPVNDKACSIKGSFRCLSSQFRGNETGLITPNKHGLQNPYRIPPLF 984
D 2322 ITTDNKVIEKENELIRDLFTQIE-----CLMSD----- 2349
QY 985 PISFVTKCKKMLLENEEE-----HMSPER----- 1012
D 2350 -----QECVKRREBEITQOLNEVIEKLOQELANIGOKTSMAHSLSEADSLKHQLDV 2402
QY 1013 -----EMGNENIPSTVSTISRNNIRENVEK-----EASSNINEVGS 1049
D 2403 VIAEKLEQOVETANEEM-----TEMKNVLKENFPMNOLQLEFLSKRERSEVETQS 2457
QY 1050 -STNEVGSSITNEIGSSDENIOAELGRN-----RGPKNAMLRGLVQ 1090
D 2458 IPNSVVAADHLSKDKPELEVLTEDALKSLNQYFKSPEENGKSIINLETRLOLE 2517
QY 1091 PEVYKOSLPSCNCKHPKIKQOEYEVQVYNTDFSPYLINDLEOPGSSHAOVOCSETP 1150
D 2518 STYSANDLEITQC-KYQIKDMQOGOFETMLQKIVNLQKIVBEKAAALVSOIQLEA- 2575
QY 1151 DDLIDGEEKEDTSFAENDIKESSAVFSKSVQ-----KGLSRSPSPPT----- 1194
D 2576 -----VOEYAKFCODNOTISSEPERTNIQNLQREDELSDSLALIRISELESQV 2627
QY 1195 ---HTHLAQYRGAKAKLESSENLSDEBELPCFOHLLFGKYNINIPSGSTRHSTVATEC 1251
D 2628 VEMHTSLI-----LEKQVEIAEKNVLEKEKLELLEKGENEKKOREKREKSSPODEV 2683
QY 1252 LSKNTEENLSTKNSLNDSCNOYLAKASQEHLSERTKSASLFSQSCSELDTLANTN 1311
D 2684 LKTTTTE---LPHSNESGFEFNELEALRAE-----SVATKAELASYEKAKERIQDELLVKE 2735
QY 1312 TODPFLIGSKOMR-HQSES-QGVGLSDKELVSDDEER----- 1347
D 2736 TMTNTSLQKDLQYRDLHALEAKELSTLEKEDTEVEGESKACMFEPPLITLSKSISQTD 2795
QY 1348 GT-GLENNQOEDOSMDSNUG---EASGCESEFVSDEDCGLSSQ-----SDLTLT 1394
D 2796 GTLKISSNOTPOLYKVMGQIQLNQSCESS-EVEEIIISOFTKEIEKMOELHAETLDM 2854
QY 1395 QOEDTMOHNLK-----LQOEAMAELEAVLE----- 1419
D 2855 ESHHISFETELKKNHYAVOALLEEGGTAKAVIQLRSKEVFSGFYNNCFSTLDCDSGDMG 2914
QY 1420 -----QHGSQPSNSYP----- 1430
D 2915 OGILYTHQOOFDIASGREGSESATDSPFKIKGLLRAVHNEGQVLSLTESPYDGED 2974
QY 1431 -----STISSSALIEDLRNEQOSTSEKAVLTSCSSS----- 1462
D 2975 HSIQOVSPEWLEBERRKAYINTISIKDLITKMOQLORAEYVDSQSHESFDMRGELLAL 3034
QY 1463 --YPISONPEGLSADKFEVADSST-----SKNKPEGVERSSPSKCPSLDRWYM 1510
D 3035 QOVFLERSVLLAARFELTALCTTDAVGLNCLBQRIQOGAVEYQAMECLOKADR--- 3091

```

```

QY 1511 HSCGSLQNNRNPYSQEBELI-----KVVDVEEOQLESQGHDLTETSYLPRODECTPYL 1564
D 3092 -----KSLSEIALNAQNNGRKITLKEESENKPSQELLEYNIQOQ-----SQML 3138
QY 1565 ESGISLSDPESDPSDRAPESARVGNISSTISALKVPOLKAESQSPAANTPTAG 1624
D 3139 EMQVELSS-----MKDRATELOE-----OLSSKMYVAEIK-SELQOTYLETLETLKAO 3186
QY 1625 YNMEE-----SVSRKPELTAEPERVAKRM 1650
D 3187 HKHLELEAFRLYEKKQKTDVEYHLINDTLASEQKSKRELQWALKEKAKLGRSEERKEEL 3246
QY 1651 SMVSGLTPEEFMVLVYFARKHHITLNLTEFTTHVAKMTDAFVCERTL 1701
D 3247 -----EDLKFSLESQKORNLQ-NULLQOKOLNLSQOKIESQRM 3287

RESULT 10
ID CENF HUMAN STANDARD: PRT; 3210 AA.
AC P49454: Q13246; Q13171;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE CENP-F kinetochore protein (centromere protein F) (Mitosis) (AH antigen).
GN CENP-F.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Breast carcinoma;
RX MEDLINE=95379848; PubMed=7542657;
RA Liao H., Winkfein R.J., Mack G., Rattner J.B., Yen T.J.;
RT "CENP-F is a protein of the nuclear matrix that assembles onto
RT kinetochores at late G2 and is rapidly degraded after mitosis.";
RL J. Cell Biol. 130:507-518(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95379848; PubMed=7651420;
RA Zhu X., Mancini M.A., Chang K.-H., Liu C.-Y., Chen C.-F., Shan B.,
RA Jones D., Yang-Feng T.L., Lee W.-H.;
RT "Characterization of a novel 350-kilodalton nuclear phosphoprotein
RT that is specifically involved in mitotic-phase progression.";
RL Mol. Cell. Biol. 15:5017-5029(1995).
RN [3]
RP SEQUENCE OF 2194-3210 FROM N.A.
RX MEDLINE=9536446; PubMed=7612011;
RA Li O., Ke Y., Kapp J.A., Fertig N., Medsger T.A. Jr., Joshi H.C.;
RT "A novel cell-cycle-dependent 350-kDa nuclear protein: C-terminal
RT domain sufficient for nuclear localization.";
RL Biochem. Biophys. Res. Commun. 212:220-228(1995).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=95370296; PubMed=7642639;
RA Zhu X., Chang K.-H., He D., Mancini M.A., Brinkley W.R., Lee W.-H.;
RT "The C terminus of mitosis is essential for its nuclear localization,
RT centromere/kinetochore targeting, and dimerization.";
RL J. Biol. Chem. 270:19545-19550(1995).
RN [5]
RP CHARACTERIZATION.
RX MEDLINE=98437347; PubMed=9763420;
RA Chan G.K.T., Schaar B.T., Yen T.J.;
RT "Characterization of the kinetochore binding domain of CENP-E reveals
RT interactions with the kinetochore proteins CENP-F and hBUBR1.";
RL J. Cell Biol. 143:49-63(1998).
CC - FUNCTION: PROBABLY REQUIRED FOR KINETOCHORE FUNCTION, INVOLVED IN
CC CHROMOSOME SEGREGATION DURING MITOSIS. INTERACTS WITH
CC RETINOBLASTOMA PROTEIN (RB), CENP-E AND BUBR1.
CC - SUBUNIT: HOMO- OR HETERODIMER.
CC - SUBCELLULAR LOCATION: NUCLEAR MATRIX (BUT NOT IN THE NUCLEOLUS),

```



```

Db 1565 EEEGLVPSLSSS-CVDPSSSLSSLG---DSSFYRALLEQTDGMSLNLNCGVVSANOCV 1619
OY 1174 SAVFSKSVQGGELSRSPSPFTHT-----HLAOGVGRGAKKLESS----- 1212
Db 1620 DEVFCSLSQENLTKRKTTPAPAKVGEELSLCEVYKQSLKEKMEGOCIMKKEIOE 1679
OY 1213 -EENLSSSEDELPCFQHLFGKYNIPSOST-----RHSTVATECLSKNNEILLKNS. 1266
Db 1680 LEQULSSERELDCLRKQYLSBNQMOQKLTSLYLEESKLAE--KKQEOJLSLELVA 1737
OY 1267 L-----NDCNQVILAKASEHHLSEETKCSASLFPSSQCELEDLTANTPTDPPFLIGSS 1321
Db 1738 RLQLOGJLSSRSILGIDTEDAIOGRNESCDS-----KEHTSETTERP----- 1782
OY 1322 KQMRHOSQSGVGLSPKEUSD--DEER--GTGLENNOE---ECSMDNL---GE--- 1367
Db 1783 KHHVHQ-----ICDDAQODLNLDEKLTETGAVKPTGECGSEOSPDNTYPPGDKT 1835
OY 1368 -AASGESETSVS-----EDCSGLSQSDILTTQOR--DTMOHNL-----IK 1406
Db 1836 QGSSECTSELSEFGPNALVMDPLGNQEDIHNLQLRKETSNNELRLHIEDRKRYES 1895
OY 1407 LQDPALEAVLEFQHSQPSNTPSIISDSSALEDLRNPQOSTSEKAVLTSQSSSE-YPI 1465
Db 1896 LLEMEELDSKHLQEVQYLTKEACIELEKIVGELKKNESDLSEKLEYFSCDHOELQR 1955
OY 1466 SONPEGLSADKPFVSDSTSKNKEPQVERSSPKCPSLDRLWYMHSCSGSLNRRNPQ 1525
Db 1956 VERSEGLNSD-LEMHADKSSREDIGDNVAK-----VDSW-----KERFLDVE 1997
OY 1526 EELIKV---VDVEEQULESGPHDLTETSYPRLDLEGTPLYLESGLFSDPESPSE 1581
Db 1998 NELSRIRSEKASLEHVALYLEADLEVYQTEKL---CLEKDNENKQKIVYLEELSVYTS 2054
OY 1582 DRAPESKRVNINSSSALKVPOLYKVAESQSPAAHTTPTAGYNAMESSVSEKPEL 1639
Db 2055 ERNQLRGEIDTMSKRTTALDQLESEKMEKETOE--LESHQSECLHCIOVAEAVERKETEL 2111

```

## RESULT 11

YMF67\_YEAST STANDARD: PRT: 1658 AA.

```

AC 003661; Q04988;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE Hypothetical 187.1 kDa protein in GOAL-Erg8 intergenic region.
GN YMR219W OR YMR261.13 OR YMR959.01.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN RA
RP SEQUENCE OF 1-711 FROM N.A.
RC STRAIN=5288C / AB972;
RA Dedman K., Brown D., Bowman S., Barrell B.G., Rajandream M.A.,
RA Walsh S.V.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN RA
RP SEQUENCE OF 608-1648 FROM N.A.
RC STRAIN=5288C / AB972;
RA Stelton J., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z49309; CAA8934.1; -

```

```

DR EMBL; Z49309; CAA8934.1; -
DR SGD; S0004832; YMR219W.
KW Hypothetical protein.
SQ SEQUENCE 1658 AA; 187137 MW; 3893F968305A757D CRC64;

```

## Query Match

3.3%; Score 322; DB 1; Length 1658;

Best local similarity 19.7%; Pred. No. 2.8e-06;

Matches 344; Conservative 255; Mismatches 704; Indels 444; Gaps 76;

```

OY 132 NNAKRLQSEPNPSLOETLSVQLNLTGVRILTKRQRIQPOKTSYVIELGSDSDTV 191
Db 97 NRKSLITLTLEKDALFERSLSEEE-----RQRQ-----LHDSLM 131
OY 192 NKATYCVSGDQELQI--TPQGT-----RDEISLDAKKAACEFSDVTNTHNQPSN 243
Db 132 NKYTGNKSKHQRILDLAKSQTGTDTSQNNDELPLDS-----FISSPLDADESSN 184
OY 244 NDINTTEKRAEHRPEKYQGSVSNLHVPCGTNTNTHASLOHENS-----SLLTQR- 296
Db 185 IDSQKDEDLQKQ-----SLIKDPLE--NDEVELSEKKNKSDQSSPSIMILSDEE 234
OY 297 -----MAYEKAFCNKSQPGIARSOHN-----RMAGSKETCDRRTPSTEEK 339
Db 235 YAEAGALQDVSNDEYAEQOVERKNIGQOAVENATQISSDSSQGVYSEGVEMELE 294
OY 340 VDLNADPLCEKKE-----NNKQKLPCSENPRTQEDVPMI--TLNSSIQKVMWF 386
Db 295 DDIDVESDAKDESGQAGEGTEHVSDFSKYMQPRTDNTK---IPVIEKYSDEKHVQRY 350
OY 387 SR-----SDELLGSDSDHSGESNAKVADYLDVLENDVEYSGSSEKI--DLASDPHE 438
Db 351 SEDGAFDQFSGVNIYVDESEDEESQASYSANMENYVHNHEHLEDKELLEDESSD--- 407
OY 439 ALICKSERVSKSVESNIEKIGKTYRKASLDPNLSHVTEMLIGAFVTEPQIIOERPL 498
Db 408 -----SESGAQSQSEQSEDDEFEKMKNEKSTSEETNTSES-----RDQGRKADAYT 455
OY 499 TNKLKRRRPTSGHLRFDIKKADLAQKTPPMINQGTNOEDQGVNMTNTHSGHEKTK 558
Db 456 KMKVEQDENE--EPE---KODILRSSLDKFNHNNKSEYSENVL-----ENETD 501
OY 559 GDSIQNEKNPNPIESLEKESAFKTKAPRISSTISNMLELNIH----- 601
Db 502 PALVERENQINDVQGV-----VTGKVEEDLHHPSPDNLYDLAARMQ 546
OY 602 -----NSAKPKKRLRRKSSRTHIHALELVSRNLSPNCTEQLDSCSSSEETKKKYN 656
Db 547 FOOSRNSNCPQKE-----EQVSESYLGHNSGNSNLGRLDSESE----- 585
OY 657 QMAYR-----HSRNLQMECKRPATGAKKSNKPNQDTSKRHDSDTPPELKLITNAPGSFTK 711
Db 586 QIPKADPTGNNNNKLTDRGDLSSVIEVEKVE---KRLDSTKEKL---VPLSTDT 638
OY 712 CSNTSEL--KEFNPSLPREK--EEKLETQKVSNNADPMDLMSGERVLQTERSVSSS 768
Db 639 TINNSLGNEDSIYSLDADALSENLTDPIMEIKTTPVEYVYSSV--YSSSYEDNT 697
OY 769 ISLVP-----GTDYGTQESISILEVSTLCKAKTEPKVCVQAFAENPK 812
Db 698 VAMPPOVEYTPPMNDPFNSLNDYERKHDLLKSTIALAALPAFTKQD-----AEFVE-A 750
OY 813 GLIHGSKQRNDTEGKKYVLLGHEVHNSRETSTIEMESELDQX--LQNTFKVSKRS--F 869
Db 751 GVTKSLTTSYGHNTIF-----HTSKETKQVS--DLDESTNVTVFENNTDKNKNSKNF 804
OY 870 ALFSNPGMADEECATSAHSGSLKQSPKVTPECEQKEENQAGNESNI---RPVOTVNI 925
Db 805 PGVANSIDKSTEDNTEDEKYSALNTYNTVGDSSCEDIIETASVWEELRCEKMDMAEM 864
OY 926 TACFPVVGQKDKVYDNAKSKIKGSRPLSS--QFR--GNETGLITPKHGLLQNPYRIPP 982
Db 865 SSGDECVKQND--DGSKTQIS---FSTDSPNQFQESNDNTFFSSTKYKV-----RNSD 912

```

```

QY 983 LFPKSPVKTCKCKNML-----EENFEHSMSP-----REKNGENI-----PSTVSTI 1026
D 913 LEDDESLAKELTAKAEVOKLDEDESEDSEYEOYADPEPNDGESSNIKGTCKKDLGIV 972
QY 1027 SRNRNR-----ENVEKASSNNI-----EVG-----SSTNEY 1054
D 973 EPENEKVKVHEEETLFEANVSSSVVQKDKMHTDVIINOEAQNYEAGERKYITQNTDTE 1032
QY 1055 GSSINEIGSSDENIQAELGRNCPKLNMLRLGVLOPEVYKOSLPSNCKHPEIKKOYE 1114
D 1033 EAHITIERIDEN-----AIGNN-----MEIPRSCEVETHNEVLE 1069
QY 1115 EVVOVTV-NTDPSPLYISDLNLEOPMGSSHAQVCEPDDLLDGEIKEDTFAEN-DIK- 1171
D 1070 RRATTEIEMTK-----ALENNITNM-----HDQVQACSDSDRQDSTAEKNEVGSAKHNDIRY 1122
QY 1172 ESSANFMSKVQKGLSRSPPTHTHLAGYRGAKKLESSENL-----SSEDE 1221
D 1123 SSSSEIESVEPLKPESDRSNIFFSSPIRVIGAVVGVGVAVVAESFYKIDVMDSESDDVV 1182
QY 1222 ELPCFOHLFGKVNNIIPSOSTRHSTVATECLSKNTEENILS--LKNSLDCSN--QVI-- 1275
D 1183 DIGOVNODIFKNSNTDASVNMKSVSKRDEDEDAVILGVTAAHNDNGNSRVINI 1242
QY 1276 -----LAKASOEHHLSEETKCSASLFSSQCSLEDLTANTQDPLI 1318
D 1243 DPTNGAYEEDSEVFRQGVCKENLHKSEEPVLE-GLQSEQHFEKKHSENEEFPDIYG 1301
QY 1319 GSKKMRHROSESQV-----GLSDKELYS-----DDEERGIGLENNOEBOGMSDNL-- 1365
D 1302 DITSANIHSNAPDDIKROQLLNLSLDLENYSQLIEDSRGKNKOESESDVNTSREDDLEF 1361
QY 1366 -----GEASGCESETSVSEDCSLSSQS-----DILTQOQDRTQOHLNLIKQOEMAL 1414
D 1362 EKVNEKAKAGALIEEDTFSE--LDISIQHHEHEDDLNNOGRSTIELNSEPEALYEL 1419
QY 1415 EAVLE-----OHGSOPSNYS--IISDSSALE--LNPQOSTSEKAV 1454
D 1420 E--IEGPTEAASSKMNDEDERGNIPTDLPSPDSPEEYTDSPYSENSEITAEKSA 1477
QY 1455 LTSOKSEYPIQONPEGLSADKEFVSADSTSKNKEPGV-----EBSPSKCPSLDDMY 1509
D 1478 PTPSPVYEL-PSDTPNEVPMELNDELPAITLKHDKTNTSVLDDSEHLSSHVDVNEPH 1536
QY 1510 MHSQSGSLQNRNYPQOELIKVADVEEQULESGPHDLETSTYLPRODLEGTPYLESGIS 1569
D 1537 DNSINIKVNEGEEPEHQAVIDIPKVEVKEQEEEMP-----SKSVLEEQ-----KPSME 1584
QY 1570 LFSD--DESDPSED-----RAPESARVGNIFSSTSALKVVPOLKVAESQSPAANT 1619
D 1585 LINDKSSPENNDDETREKDKTKAKKRSRKNYNSRRRRKRTTEGSSAASNTKRRRGHE 1644
QY 1620 TDTAGYN 1626
D 1645 PKSRGN 1651

```

```

RESULT 12
ANK3_HUMAN
ID ANK3_HUMAN STANDARD; PRT: 4377 AA.
AC Q12955;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ankyrin 3 (ANK-3) (Ankyrin G).
GN ANK3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain stem;

```

```

RX MEDLINE=95138209; PubMed=7836469;
RA Kordeli E., Lambert S., Bennett V.;
RT "Ankyrin. A new ankyrin gene with neural-specific isoforms localized
RT at the axonal initial segment and node of Ranvier.";
RL J. Biol. Chem. 270:2352-2359(1995).
CC -1- FUNCTION: Membrane-cytoskeleton linker.
CC -1- ALTERNATIVE PRODUCTS: A number of isoforms are produced by
CC alternative splicing.
CC -1- TISSUE SPECIFICITY: Expressed in brain and other tissues.
CC -1- SIMILARITY: CONTAINS 23 ANK REPEATS.
CC -1- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U13616; AAA64834.1; -
DR HSSP: P55273; 1B18.
DR Genew: HCN0494; ANK3.
DR MIM: 600465; -
DR InterPro: IPR002110; ANK.
DR InterPro: IPR000488; Death.
DR InterPro: IPR000906; 205.
DR Pfam: PF000531; ank; 24.
DR Pfam: PF00791; death; 1.
DR PRINTS: PRO1415; ANKYRIN.
DR SMART: SM00248; ANK; 21.
DR SMART: SM00218; 205; 1.
DR PROSITE: PS50088; ANK_REPEAT; 21.
DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE: PS50017; DEATH_DOMAIN; 1.
KW Cytoskeleton; Alternative splicing; Repeat; ANK repeat.
FT REPEAT 73 102 ANK 1.
FT REPEAT 106 135 ANK 2.
FT REPEAT 139 168 ANK 3.
FT REPEAT 172 201 ANK 4.
FT REPEAT 203 230 ANK 5.
FT REPEAT 234 263 ANK 6.
FT REPEAT 267 296 ANK 7.
FT REPEAT 300 329 ANK 8.
FT REPEAT 333 362 ANK 9.
FT REPEAT 366 395 ANK 10.
FT REPEAT 399 428 ANK 11.
FT REPEAT 432 461 ANK 12.
FT REPEAT 465 494 ANK 13.
FT REPEAT 498 527 ANK 14.
FT REPEAT 531 560 ANK 15.
FT REPEAT 564 593 ANK 16.
FT REPEAT 597 626 ANK 17.
FT REPEAT 630 659 ANK 18.
FT REPEAT 663 692 ANK 19.
FT REPEAT 696 725 ANK 20.
FT REPEAT 729 758 ANK 21.
FT REPEAT 762 791 ANK 22.
FT REPEAT 795 825 ANK 23.
FT REPEAT 828 857 ANK 24.
FT DOMAIN 1519 1898 SER-RICH.
FT DOMAIN 4090 4174 DEATH.
SQ SEQUENCE 4377 AA; 480399 MW; F42379E55768B684 CRC64;

```

```

Query Match 3.3%; Score 321.5; DB 1; Length 4377;
Best Local Similarity 18.1%; Pred. No. 1e-05;
Matches 403; Conservative 352; Mismatches 858; Indels 613; Gaps 98;

```

```

QY 6 LRVEEQVNIYNNKQKILBPI-----CLELIKPEVSTKCDHLFCFKFCKMLLNKQKGPQ- 60
D 2238 MRVKEETHITTTTRRVYHSPGEGASRIETMSV-----HDIKKAFQSGRDPQKE 2289

```

QY 61 -----CPLCKNDITKRSLOESTRISOLVEELIKITICAFOLDTGLEVANSYFAK 109  
Db 2290 LAGLFEHKSAAVSPVHKSAAETSAOHAERKQNMKPKLEIRI-EVHIEKQNAEPTLEVIIR 2348  
QY 110 KENNSPE-----HLKD-----EVSIIOGMYRNRAKRLLOSPERPISOETLSVOLSLUG 160  
Db 2349 ETKHNPKEKMYVYOKDLISRGDINIKDLPEKHDA---FPCSEGOQOEELTAEESLPS 2405  
QY 161 TVRTLRTRKRIOPORTS--VYIELGSSSEDTYVANKATYCSV--GDOELLQTPGTRDEI 216  
Db 2406 YLESSRWNTVPYSQEDSRPSSAQLISDDSYTKLLLSOHSIEYHDELDELSELGSSYRAE 2465  
QY 217 SLSAKRAACESESTDYVTEHOPSNNDLTTEKRAER--HPEK-----YOGSS-- 265  
Db 2466 KMLSEKLDVSHSTEEVTFDHAGPSSSELQSDKRSREKATAPKKEILISKIKYDSEN 2525  
QY 266 -----VSNIHAE--PCGTNTHASLOHENSLS-----LTKDMN--VEKAE 303  
Db 2526 GVGKVSXDEHFDKVTYLVHSGNVSSPKHAMMRTEDRLDGRKELIYEDRVDTYKAE 2585  
QY 304 -----FCNKSQOPG---LARSQHNRMAGSKETCNDRTPTTEKKVVDLNAPLCERK 351  
Db 2586 EKLTEVSGFFRDKTEKLNDELQSPERKARPKNGKEYSSQSTSSPERV-LTTELASND 2644  
QY 352 EWNK--OKLPCSEN--PRDTEDPWITLNNSTIOKYNEWFSRDELLGSDSDHSGSESNAK 408  
Db 2645 EWMKARQHPDQGGFPAKEEKAP--SLPSSPEK-----WVLSQTEDEKSTYWEAK 2692  
QY 409 VADVLVDVINEVDEKSGSEKIDLLASDPHEALICKSEVHKSVEYNEDK----- 459  
Db 2693 GS-----ISQKADBDGPOGFOLOKQSLKSLIRLKEFGTAKSKDMQSDKSDGOSRIP 2747  
QY 460 -----IFGTYRKKA--SLPULS-----HYTENIILAFYTER 490  
Db 2748 VKTIOESKLPYQYFARKKQOKAIDLDEDESYSVOKDEMYLTKQEHQOSNEI----- 2799  
QY 491 QIOERPLTYLKKRRRTSGLHPEDEIKK--ADLAOVKTPEMINOG--TQOTONGQVM 546  
Db 2800 -VVNDGSDVNYKKQRTKMSKAMPDSFSEQOKDLACHITSDLATRGWMDKVKVTRTMS 2858  
QY 547 NITSGHENTKIGSIONEKNPNPI-----ESLEKESAFKKAEPDISSISNMEL-ELNT 600  
Db 2859 GATNNKSOKEKLSHVLDVHRENIIGHPEKSVDOKEFMVTERERKLLNGSLSEIKE 2918  
QY 601 HNSKAPKKNRLRRKSTFRHIALFLVYVRNLSPNCTELOIDSSSSSEIKKKKYNOVP 660  
Db 2919 MTVKSPSKKTYLR-----EYVYKEGDHPGGLD-----OPSRSESSSAVSHIPV 2962  
QY 661 RHSRNLQMEGKEPATGAKKSNKPNEDQTSKRHSDTEPE-----LKLTHNAGSFPTK 712  
Db 2963 RVADERRMLSSNIDGFCQSAFPKHLSOKLSOSSMSKEVETQHNSIDEKAVYSEI 3022  
QY 713 SNTSELKEFVNPISLPREKE-----EKLETYKVSNNADDPKDLMSGERV-----LQ 759  
Db 3023 SKVSKHOSYGLCPLEETETSPTKSPDSLSEFSGKSPSSDVDPHSPIGOLEKLAPLQ 3082  
QY 760 TERSESSSISLVGTQOESISLLEVSTLGAKTER--NKCYSO-CAAFENPGLI 815  
Db 3083 TEGSEKIKTLPIVY-----VSFVQVKQYKEKEIIOGQGVKKIISOECTVQETGTF 3132  
QY 816 H-----GSKDNRNDETEG---KYPLGHEVHNSRSTSIEMESELEDAQYLO 858  
Db 3133 YTRHQOKOPSPQSPDDTLQVYSLDSSCKSPLETPSPSESVSEVSPISKTJDSLL- 3190  
QY 859 NTFEVKSRQSALEFNSNGNAEEECATSAHSGSLKQSPKVTCECQKLEENOGKNESNIK 918  
Db 3191 -AYIPGKPSPIPEVSESEEEQ-----AKSTSLKQTTVEET-AVEREMPDVSKDSNOR 3243  
QY 919 PVQTVNITJAGPVVQOKDPVNDNAKCSIKGSRCLCSOPRGN-----EGLTIPN 969  
Db 3244 PKNNRVAVIERP---PPPLD-----ADQIESDKKHHYLPKEKVDMEVN 3285  
QY 970 -----KHGL-----LQNPYRIPP-----LFPITKSFV----- 990

Db 3286 LODEHDKQOLAEPIVIRQPPSPVPBGADVSDSSDESIYQVPYKKYFKLEVEDEQKE 3345  
QY 991 --KTKCKKNLLEENFEHSMSPEREMGNENIPVYVSTISRNNINENYKESSSNINEVG 1048  
Db 3346 KPKASAEKASQKLEESNGSKDNEFG-LGDSFQNEIAGQNNQDOSTTECSIATFAEFS 3404  
QY 1049 SSTNEVG--SSINEIGSSDEN-----IOAEIGRNG-----PKL 1080  
Db 3405 HDTATLELSDJGDVLDQDEDDGLTESDSKLEIQAMEIKKQDWNTEGLKPADRFSQSKL 3464  
QY 1081 NAMLRGLVLOPEVYKQSLPGSNC-----KHPEIK----- 1109  
Db 3465 EVIEEKGVPDEDEKPPSKSSSEKTPDKTDQSGAQEFFTEGHRPDRSVPPDYFSYKV 3524  
QY 1110 KOEVEVQVQNT---DESPYLISDNL-----EQPGSSHAQVCSFETDD 1152  
Db 3525 DEEFATPRTVATGLDFPW--SNNRGDVEVFDKSRDETKRPGGLAVEDRSPATYP-- 3580  
QY 1153 LLDGELIEDTSFAENDIKESSAVFSKQGLSRSPPTHTHLAGYRRGAKLESS 1212  
Db 3581 -----DTPPARTIDESTPTSEP-----NFPFHEKMEFMTSGA--IDMS 3620  
QY 1213 EENLSEDELEPCFO--HLLFGK-----VNNIPSOSTRHSTVATECLSKNTEE 1238  
Db 3621 KRDP--VEERLOFQIGHTSEKSGDQEGDKSMVATPQPOGDDTVEIN-LEARNET 3677  
QY 1259 NLSLSKSL---NDCSNQVILAKASOEHLHSEFTKCSLSFSSQSELELTANTNODP 1315  
Db 3678 PVEPNPSPITPSGEC-----QEGTSSGSLKESAAA-----TMTSKVP 3716  
QY 1316 FL-----IG--SSKOMRHOSESOQVGLSDK-----ELVSDDEBGTGL- 1351  
Db 3717 KLRTPIKMGISASATYTKKKEGPG-ETIDKLAVMPTSCOGLENETITMNSTANSOMGYRP 3775  
QY 1352 -EENNOEBSMDNMGEAASGESETSVEDCSGLSSQSDILTYOQRDTMOHNLKL--- 1407  
Db 3776 HEKHDFQKDNFNNNNNLDDSTIQTDNIMSNIV-LTESAVTCTEKXD---NPVKVSSG 3829  
QY 1408 -----QEMAELEAVLEQ-----HGSQPSNSPISISDSALEDL 1442  
Db 3830 KKTGVLQHCVRDQKQVLEQOQKTELGIQOKKLPKATKSPDTPPNHMSMTKASKM 3889  
QY 1443 RNPOSTSEKAVLTS---QKSEYPIQONPEGLSADKREVSADSTSKNKP--GVERS 1497  
Db 3890 KOVSQSEKTKALTTSSCYDVAKSRIPVKTPRD---NIIAVKACATQKQOGPERKAKOL 3946  
QY 1498 PSKCP-SLIDRWYHSCGSLQNNRNPQOEDLIVVDVEEQULESGPHDL-----E 1549  
Db 3947 PSKLPVKVRSVCVTTTTATTTTTTTTTTSCVVKYKSQLKEVCHSIEYFGISGE 4006  
QY 1550 TSYLPRODLETPYLEGSIISFDDPPSDPSEDAPASARVGNIPSSIALKVLQVYAE 1609  
Db 4007 TLKLVDRLESEKQMOSEL---SDEESTSRNTLSIETSR--GGOPSVYT-----KSAR 4055  
QY 1610 SAQSPAAHNTDTGYVNAEMESVSRKPELTASTERVNKRMSVY----- 1654  
Db 4056 DKKTEAPLAKSKSEKASSEKSSRRTPQ--SPEERDINATVADHLGLSWTLAEELN 4113  
QY 1655 -----SGLTPEFMYLFARKHNTLTLTTEETHYVMK-----TDAEFVC 1697  
Db 4114 FSVDEINOIREVNNLSIOSFMFLKKVYIRDG--KNATTDALTSVLTKNRIDYITLL 4170  
QY 1698 ERTIKTYFLGIAGGVVYVYFWTOSIERKMLNHEDEVRQDVVYGNRHHQPKARBS-- 1755  
Db 4171 EGPITFDGNISG-----TRSFADEN-----NVPHDPVDQMNQENSSCNLESCEA 4213  
QY 1756 QDRKIFRGLIECCYGPPTNMPDQLEWVYOLCGASVYK-----ELSSFTLGTVPIVYQ 1811  
Db 4214 QARRVITGL-----LDRLDSPDQCRDSITISYKLGEGKFE-ANGSH--TEIT 4258  
QY 1812 PDAMTE 1817  
I I I I I

DB 4259 PEAKTR 4264

RESULT 13

ATRX\_HUMAN STANDARD: PRT: 2492 AA.

AC P46100; P51068; Q15886; Q9NTS3; Q9H0Z1;

DT 01-NOV-1995 (Rel. 32, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Transcriptional regulator ATRX (X-linked helicase II) (X-linked nuclear protein) (XNP) (ZNF-HX).

GN ATRX OR RAD54L OR XH2.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3; 4 AND 5), VARIANT S-1860, AND VARIANTS ATR-X.

RX MEDLINE-97123494; PubMed-8968741;

RT Picketts D.J., Higgs D.R., Bachoo S., Blake D.J., Quarrell O.W.J., Gibbons R.J.;

RL "ATRX encodes a novel member of the SNF2 family of proteins: mutations point to a common mechanism underlying the ATR-X syndrome.";

RN Hum. Mol. Genet. 5:1899-1907(1996).

RN [2]

RP SEQUENCE FROM N.A. (ISOFORMS 2 AND 4).

RX MEDLINE-97386582; PubMed-9244431;

RT Villard L., Lossi A.-M., Cardoso C., Proud V., Chiatoni P., Colleaux L., Schwartz C., Fontes M.;

RL "Determination of the genomic structure of the XNP/ATRX gene encoding a potential zinc finger helicase.";

RN Genomics 43:149-155(1997).

RN [3]

RP SEQUENCE OF 860-2492 FROM N.A.

RX MEDLINE-95179111; PubMed-7874112;

RT Stayton C.L., Dabovic B., Gulisano M., Gecz J., Broccoli V., Giovannazzi S., Bossolasco M., Monaco L., Rastan S., Boncinelli E., Bianchi M.E., Gonzalez G.G.;

RL "Cloning and characterization of a new human Xq13 gene, encoding a putative helicase.";

RN Hum. Mol. Genet. 3:1957-1964(1994).

RN [4]

RP PRELIMINARY PARTIAL SEQUENCE FROM N.A.

RX MEDLINE-94214473; PubMed-8162050;

RT Gecz J., Pollard H., Gonzalez G., Villard L., Stayton C.L., Millassau P., Khrestchatsky M., Fontes M.;

RL "Cloning and expression of the murine homologue of a putative human X-linked nuclear protein gene closely linked to PGK1 in Xq13.3.";

RN Hum. Mol. Genet. 3:39-44(1994).

RN [5]

RP SEQUENCE OF 2401-2492 FROM N.A., AND VARIANTS ATR-X.

RX MEDLINE-95211835; PubMed-7697714;

RT Gibbons R.J., Picketts D.J., Villard L., Higgs D.R.;

RL "Mutations in a putative global transcriptional regulator cause X-linked mental retardation with alpha-thalassemia (ATR-X syndrome).";

RN Cell 80:837-845(1995).

RN [6]

RP SEQUENCE OF 1375-2492 FROM N.A.

RX Pearce A., Chapman J.;

RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

RN [7]

RP MEDLINE-98167853; PubMed-9499421;

RX Cardoso C., Timst S., Villard L., Khrestchatsky M., Fontes M., Colleaux L.;

RL "Specific interaction between the XNP/ATR-X gene product and the SET domain of the human EZH2 protein.";

RN Hum. Mol. Genet. 7:679-684(1998).

RN [8]

RP SUBCELLULAR LOCATION, AND ASSOCIATION WITH PERICENTROMERIC HETEROCHROMATIN.

RX MEDLINE-20040663; PubMed-10570185;

RA McDowell T.L., Gibbons R.J., Sutherland H., O'Rourke D.M., Bickmore W.A., Pombo A., Turley H., Gatter K., Picketts D.J., Buckle V.J., Chapman L., Rhodes D., Higgs D.R.;

RT "Localization of a putative transcriptional regulator (ATRX) at pericentromeric heterochromatin and the short arms of acrocentric chromosomes.";

RL Proc. Natl. Acad. Sci. U.S.A. 96:13983-13988(1999).

RN [9]

RP DISEASE.

RX MEDLINE-20213147; PubMed-10751095;

RA Villard L., Fontes M., Ades L.C., Gecz J.;

RT "Identification of a mutation in the XNP/ATR-X gene in a family reported as Smith-Fineman-Myers syndrome.";

RL Am. J. Med. Genet. 91:83-85(2000).

RN [10]

RP VARIANT ATR-X SER-1713.

RX MEDLINE-97196774; PubMed-9043863;

RA Villard L., Lacombe D., Fontes M.;

RT "A point mutation in the XNP gene, associated with an ATR-X phenotype without alpha-thalassemia.";

RL Eur. J. Hum. Genet. 4:316-320(1996).

RN [11]

RP VARIANT JM GLN-2131.

RX MEDLINE-96224392; PubMed-8630485;

RA Villard L., Gecz J., Mattei J.-F., Fontes M., Saugier-Verber P., Munnich A., Lyonnet S.;

RT "XNP mutation in a large family with Juberger-Marsidi syndrome.";

RL Nat. Genet. 12:359-360(1996).

RN [12]

RP VARIANTS ATR-X.

RX MEDLINE-97467722; PubMed-9326931;

RA Gibbons R.J., Bachoo S., Picketts D.J., Affinos S., Aseunauer B., Bergoffen J., Berry S.A., Dahl N., Fryer A., Keppler K., Kurosawa K., Levin M.L., Masuno M., Neri G., Pierpont M.E., Slaney S.F., Higgs D.R.;

RL "Mutations in transcriptional regulator ATRX establish the functional significance of a PHD-like domain.";

RN Nat. Genet. 17:146-148(1997).

RN [13]

RP VARIANT ATR-X LEU-246.

RX MEDLINE-20123062; PubMed-10660327;

RA Fichera M., Romano C., Castiglia L., Falla P., Ruberto C., Amata S., Greco D., Cardoso C., Fontes M., Ragusa A.;

RT "New mutations in XNP/ATR-X gene: a further contribution to genotype/phenotype relationship in ATR/X syndrome.";

RL Hum. Mutat. 12:214-214(1998).

RN [14]

RP VARIANT SHS LYS-1742.

RX MEDLINE-99347960; PubMed-10417298;

RA Lossi A.-M., Millan J.M., Villard L., Orellana C., Cardoso C., Prieto F., Fontes M., Martinez F.;

RT "Mutation of the XNP/ATR-X gene in a family with severe mental retardation, spastic paraplegia and skewed pattern of X inactivation demonstration that the mutation is involved in the inactivation bias.";

RL Am. J. Hum. Genet. 65:558-562(1999).

RN [15]

RP VARIANT CMS THR-2050.

RX MEDLINE-99326061; PubMed-10398237;

RA Abidi F., Schwartz C.E., Carpenter N.J., Villard L., Fontes M., Curtis M.;

RT "Carpenter-Waziri syndrome results from a mutation in XNP.";

RL Am. J. Med. Genet. 85:249-251(1999).

RN [16]

RP VARIANTS ATR-X E-175; 178-V--K-198 DEL; S-190; P-219; L-246 AND C-249.

RX MEDLINE-99219535; PubMed-10204841;

RA Villard L., Bonino M.-C., Abidi F., Ragusa A., Beloungue J., Lossi A.-M., Seaver L., Bonnefont J.-P., Romano C., Fichera M., Lacombe D., Hanauer A., Philipp N., Schwartz C.E., Fontes M.;

RT "Evaluation of a mutation screening strategy for sporadic cases of ATR-X syndrome.";

RL J. Med. Genet. 36:183-186(1999).





```

Db 1033 KQKNGTDDGKSKKIRKDTSKKKDELSDYAEKSTGKDGSDSDSKNGAYGREKK 1092
OY 1256 TEENLTSLKNSINDCSNOVILAKASOEHHLSEETKCSA--LFSSOCSELEDTFANTNT 1312
Db 1093 RCKLLGKSKRRKDDCS-----SDPEKTSMEKEDGNSSSDKRLKRIELERRRLSSKRT 1146
OY 1313 QDPFLIGSSKQMRHOSE---SOGVGLSDKELVSDDEERGTLSENNQEQSDMSNLGEA 1368
Db 1147 KEIQSGSSSSDAEESSEDNKKKQRTSSKKKAVIYEKKRNSLRTSTKRQADITSSSSS 1206
OY 1369 ASGCSEFHSVSDSCGSLSSQSDILTTQQDNTQMHNLKIQOEALEAVLEQSGSP-SN 1427
Db 1207 DIEDDDQNSIGEGSS-----DEQKIKPYTEMLV-----LSHTGFCQSSGDALSK 1252
OY 1428 SYPSTLSDSALEDLNPPOSTSEKAVLTSOKSESEPISONPGLADKFEVADSTSK 1487
Db 1253 SYPVYVDD--DDNDPENRIAKKMLLEIKAN--LSDEGSSSDDEEESKRTGCK 1306
OY 1488 NKE-PGVER-----SSPSKCPSLDRMYHSCSGSLQNRNPSQEBLLIKVVDY 1534
Db 1307 NEENPGDEFAKKNQVNSESDSEESKPPHYRHLRLHK-----LTVSDG 1350
OY 1535 EEOQLEESPHDLTETSTYPR-----QDLEGTPYLESGIS--LFSDPESDPEEDAPES 1587
Db 1331 ESGEKKTTPKEHKEVKGRRNRKVSSESDSDFOESGVSEEVSESEDQRPRTREAKKA 1410
OY 1588 ARVGNIPSTSLAKVPOLKVAESAQSPAAPHTTDTAGYNAMEEVSREKPE 1638
Db 1411 ELEENORSYKOKKKRRRIKVOEDSSSEKNSNSEFEKEKEEKEEKEEKEE 1461

RESULT 14
P531_HUMAN
ID P531_HUMAN STANDARD; PRT: 1972 AA.
AC Q12888;
DT 15-JUL-1998 (Rel. 36, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Tumor suppressor p53-binding protein 1 (p53-binding protein 1)
GN (53BP1).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N. A., AND SUBCELLULAR LOCATION.
RC TISSUE=Skeletal muscle;
RX MEDLINE=94421332; PubMed=9748285;
RA Iwabuchi K., Li B., Massa H.F., Trask B.J., Date T., Fields S.;
RT "Stimulation of p53-mediated transcriptional activation by the
RT p53-binding proteins, 53BP1 and 53BP2."
RL J. Biol. Chem. 273:26061-26068(1998).
RN [2]
RP SEQUENCE OF 946-1972 FROM N.A.
RX MEDLINE=94286584; PubMed=8016121;
RA Iwabuchi K., Bartel P.L., Li B., Marracchino R., Fields S.;
RT "Two cellular proteins that bind to wild-type but not mutant p53."
RL Proc. Natl. Acad. Sci. U.S.A. 91:6098-6102(1994).
CC -!- FUNCTION: ENHANCES P53-MEDIATED TRANSCRIPTIONAL ACTIVATION.
CC -!- SUBUNIT: BINDS TO THE CENTRAL DOMAIN OF P53.
CC -!- SUBCELLULAR LOCATION: NUCLEAR. BOTH NUCLEAR AND CYTOPLASMIC
CC IN SOME CELLS.
CC -!- SIMILARITY: CONTAINS 2 BRC1 DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

DR EMBL: AF078776; AAC62018.1;
DR EMBL: U09477; AAA21596.1;
DR Genew: HGNC:11999; TP53BP1.
DR MIM: 605230;
DR InterPro: IPR001357; BRC1.
DR Pfam: PF00533; BRC1.
DR SMART: SM00292; BRC1.
DR PROSITE: PS50172; BRC1.
KW Nuclear protein; Transcription regulation; Activator; Repeat.
FT DOMAIN 1724 1848 BRC1.
FT FT 1864 1964 BRC2.
FT FT 1642 1646 POLY-SER.
FT DOMAIN 1760 1764 POLY-GLU.
SQ SEQUENCE 1972 AA; 213573 MW; 13E2C8A265FD2A CRC64;

Query Match 3.3%; Score 317.5; DB 1; Length 1972;
Best Local Similarity 17.3%; Pred. No. 5.4e-06;
Matches 389; Conservative 279; Mismatches 749; Indels 831; Gaps 88;

OY 111 ENNSPEH--LKDEVSIIQSMGYRNRKAKRLLOSEPENP----- 145
Db 23 EDSQPEQVLEDDSGSHFMSLRHLPN--LQHKENPVLDVYNSNEPOTAGEERDGNISGF 80
OY 146 --SLQETLS--VQSLNIGTVFLR--TKRIQPKTSVYIELGSDSS----- 187
Db 81 NEHLKENKVAADPVDSNLDTCGISQVLEQLPQPRRTSVYLGMSVESAFAVEEKEGELE 140
OY 188 -----EDTVAKATYCSVG-----DQELLQITPQGTREISIDS 220
Db 141 QKEKEKEEDTSCNTHH--SLGAEDTASQGLFVLELSQSDQDEVENTVYEVKDKQLQSVT 199
OY 221 AKKACAFSEETDV--TNTHEHQPSNNDLNTTEKRAE-----RHPEYQGSVSNLHVEPCG 275
Db 200 TNSGYTRLSDVQANAIKHEQGSNEDIPAEQSSKDIDVTAAQPSK-----DVHVV--- 249
OY 276 THTHASSLOHENSILLTKDRMNVKAEFCMSKQPGIARSOHNMAQSKCTCDNRTPS 335
Db 250 -----KEQMPPARSEDMPFSPKASVA-----A 272
OY 336 TEKKVDLNDPLCEKRENNKOKLPCSENPRTQEDPWTLSNLSIQKVMESRSDELIGS 395
Db 273 MEKKEQLSAQELME--SGIQIQSPPEVLSIOEDL-----EDQSKTVSS 316
OY 396 DD--SHDGESESNAKVADVLVLENDVEYSG-----SSEKIDLLASDPH----- 437
Db 317 DGCSIPSRREGGCSLASIPATTLHLQLSGQSLVQDSLSTNSSDVLAVRSPDARSTPFI 376
OY 438 -BALICKSERVHSKSVESNIEDKIFGKYRK----- 468
Db 377 VPSPTQEGRODKPMDTSVLSSEEGCFPFQKLOSGEVELENNPPLLPESTVSPQASTPI 436
OY 469 -----ASLPLNSHVTEMLITGAFVTEPQIIQGR-----PLTKKLKRK 506
Db 437 SQSTVFPPGSLPISQPFQSH-----DIFIPSPLEESQNDGKKGDMHSSSLVEEC 489
OY 507 RPTSLHPEDFETFKKADLAV-----OKTPEMINQGTNOTEQNGO----- 544
Db 490 SKTSELPEPNSEDGLSLTGSCMLMSTSEYSGSPMEBSLSHRIDEDENQIODETE 549
OY 545 -----VMNTNSG-----HENKTKGDSIONEKN----- 567
Db 550 PMSVPLNSKFVPAENDSLMNPADQGEVOLQSDNDKTKGDDTDRDISILATGCKGREG 609
OY 568 -----PNPIELKESAFKTKAEPISISSIMNE--TELNTH----- 602
Db 610 TVAEVDCIDLTGSSSQAVPSP-----ATRSEALSSVLDOEEMETKEHHPEGS 659
OY 603 -----SKAKKRLRKRKSTRHIALE-----LVYSRLSPNCTE----- 638
Db 660 SGSEVEEIPETPCESQGELEKENNESVPLHLISLITGQSGCLCQKEMPKKECEANVE 719
OY 639 ---LOIDSCSS---SEIKKKKKYN--QMPVHRHSRLQLMGKEPATGAKSKNPNDQTS 689

```

Db 720 TVTISIDSPQKAILDQLELHKEQEAEMEATSSEDSSVIVDVKEPSPRVVDSCEPLEGVE 779  
 QY 690 KRHDSTPEIKLTNAPGSFTKCSNTESEKFEVNPSPLEPREKEKLETVKVNNAEDPK 748  
 Db 780 KCDSDSWEDI---APEIEPCAEENLDTRKEKSVYEGDLKSGTAEETPEVEDSSQPSL 835  
 QY 749 ----DLMLSGERVLOTERSVESSSISLVPGDYGTQESISTL-LEVSTLGKATPEBNKC 802  
 Db 836 PLVRADDPRLRLDQLOQPOQERTSNLSLIEDSKMANAKOLSDAEAKQKIGKPSAHASQSR 895  
 QY 803 SQCAA-----FENPKLIGCCSKDNNDIEGKYP-----LGH-----EYNSRETS 844  
 Db 896 CESSSTPHFTLPK-----EGDILPLTGATPLIOLHLEPKRHSSTPLG 941  
 QY 845 IEMESELDAOYLONTFKSKROSFALFSN-----PGNAEECATPSAHS 889  
 Db 942 I-----SNPESTIATSDVWSESMVETHDPILOSGKGDCAAPDVDDKLCRLKLYVS 993  
 QY 890 GSLKOSPVYTECECEKEENOGKN-----ESNIRFVOTVNTITAGFPVVGKD--KPVN 941  
 Db 994 PETEASEESLQFLKEKPATGERKNGSTAVAESVAPQKTMVSLVSCICARQENEARSEDP 1053  
 QY 942 AKCSINGSGRFLCSSQFRNGTGLITPNKHLQNTYRIPLPLPIKSEYKTKCKK----- 996  
 Db 1054 PTPPIRGNLHPPSSQ--GEEKEKLEGDHTTIRSOQOPMKPISPVKDPVSPASOKWIOG 1111  
 QY 997 -----NLLEENFEHSMSPER-----EMGNEN-----IPSTVSTIS 1027  
 Db 1112 PSSPOEAMVTVLEQKRGSRNKENPNSKALIERSONNIGIQMECSLRPEVTSANT 1171  
 QY 1028 RNNIRENVEKEASSNINEVGSSTNEVGSSINEIGSSDENIOALEGRNPKPLNMLRLG 1087  
 Db 1172 Q-----TIKNVCEQGTSTVD-----QNFCKQDATVQTERSGSKRP----- 1206  
 QY 1088 VLOPEYKOSLQSGNCKHPEIKQETEEVYQVYNDPFSYLLSDNLEDQMGSHASQVCS 1147  
 Db 1207 -----VASPQDDTE--SLHSQGEER-----DMQPPHG--HYLHRHM 1240  
 QY 1148 EPPDLDLDGELKEDTSFAENDIKESSAVESKSVQKGLSRSPSPPTHHLAOGYRGAK 1207  
 Db 1241 RT-----IREVRLTVTRYI-----TDVYYVDGTEVERK 1268  
 QY 1208 KLESSEENLSEDEELPCFOHLLFGKVNNIPOSTRHSTVATECLSKNTEENLLSKNSL 1267  
 Db 1269 VNEETEEPI-----VECCQCEFEVPSQGTGSS 1296  
 QY 1268 NDCSN-QVLLAKASQEHLSSEETKCSASFSSQCS-----ELBDLTANTNTDPELIGSS 1321  
 Db 1297 GDLGDISPSSKASSLHRTSGTSLSA-MHSGSSGKAGAPLRGKTSSTEPADFAFPSSR 1355  
 QY 1322 KQMRHOSQGVGLSKELVSDDEERGTLLENNEQSGMDSNLGAAGCSSETSVSD 1381  
 Db 1356 GGPBKLSPPKGVSOGTPTVCEBEGDAGLGRG-----GKAPVTR 1396  
 QY 1382 CGSLSSQDILTTQORDT-----MOHNL-----IKIQEAMLEAVL 1418  
 Db 1397 GGRGRGRPPSRRTGTRETAVPGPLGIEDISPNLSPDKSFSRVPRVPDSTRIDVGACA 1456  
 QY 1419 EOHGQSPMSYPSIISDSALDLRNPEDSTEKAVLITQKSEY----- 1463  
 Db 1457 LRSDSPETIFQAAGPSDGL-DASSPGNSFVGLRVAAWSSNGFYSGKITRDVAGKRY 1515  
 QY 1464 -----PISQNEGLSADFEVSADSSSTSKNEPG-----VR 1495  
 Db 1516 KLLFDGVCYDVLGKDILLCDEPLDTEVLTALSEDEY-FSAGVYKCHRESEGLYYSTIK 1574  
 QY 1496 SPSKCPSLDDRWYMH-SCSGSLQNN-----YPSQELLKVVDVEQOLEE--- 1541  
 Db 1575 ECGRK-----WYKMAVILSLQGNRLREOYGLAPYAVPPLTKAADISLNDLVEGR 1627  
 QY 1542 -----SGPHDLETSTYLRQDLEGFPYLESGI-----SLFSDPSPDSSEDRAPGS 1587  
 Db 1628 KRNSVNSPATVASSSSSTTPTRKLTESPRASMGVLSGKRKLITSEERSPAK-NGRKS 1686

QY 1588 ARV-----GNIPSTSAALKPOLKVAESAQSPAANTDTAGY----- 1625  
 Db 1687 ATYKPAVAGAGFVSPCESDNTGEPSSAL-----EQRGLPLINKTILFLCYATLLTM 1738  
 QY 1626 -NAMEESVREK-PE-LTASTERVNRKMSWVSGLPPEEFMLYKPAKHHITLTLNLTPE 1682  
 Db 1739 ATTSDKLASRSKLPDQPTSSSE-----EEFLLEIPRNKQY----- 1775  
 QY 1683 ETHVYVAKTDAEVC-----RLKTRFLGAGKRWVSYRWYNO 1721  
 Db 1776 --TESQLRAGAGYILDFEQAQCNFAYQCLLADQHCRTKRYFLCLASGIPCVSHVW-- 1831  
 QY 1722 SIKERKMLNEHDEVDGVDVNGRHHQCP-----KRRESODRKIFRGLEICCYGPTN 1774  
 Db 1832 -----HDSCHANQLONTRNLYLLPAGYSLSEORILLMDQRE-----NPFON 1871  
 QY 1775 MP-----TDQLEWVQLC-----GASVYKELSSFT-----LGTGVRPIYVQPDWATEDN 1819  
 Db 1872 LKVLVSDQOQNFLELMSILMTGAAVSKQHHSSAHNKNDIALGVDPVVTDPSC--PA 1928  
 QY 1820 GPHATQKCEAPVYTRRWVLDVVALYQC 1847  
 Db 1929 SVLKCAEALQLPVSOEWI-----QC 1950  
 RESULT 15  
 ALM1\_SCHPO STANDARD; PRT: 1727 AA.  
 ID ALM1\_SCHPO  
 AC 09UTR5; 013113; 09UTR8;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Abnormal long morphology protein 1 (Sp8).  
 GN ALM1 OR SPAC1486.04C.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RX MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgourou K., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Feltham T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabbittowitsch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Wolchert G., Aert R., Robben J., Grynolopez B.,  
 RA Welfens T., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel G., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambut R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Leleure V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
 RT "The genome sequence of Schizosaccharomyces pombe."  
 RL Nature 415:871-880(2002).  
 RN [2]  
 RP SEQUENCE OF 495-1727 FROM N.A., AND CHARACTERIZATION.  
 RC STRAIN=972;

```

RA MEDLINE=20123449; PubMed=10660053;
RX Jimenez M., Petit T., Gancedo C., Goday C.;
RT "The atm1+ gene from Schizosaccharomyces pombe encodes a coiled-coil
RL protein that associates with the medial region during mitosis.";
RN Mol. Gen. Genet. 262:921-930(2000).
RN [3]
RN SEQUENCE OF 644-834 FROM N.A.
RC STRAIN=968 b90;
RX MEDLINE=20223868; PubMed=10759889;
RX Ding D.Q., Tomita Y., Yamamoto A., Chikashige Y., Harauchi T.,
RA Hiraoka Y.;
RT "Large-scale screening of intracellular protein localization in living
RL fission yeast cells by the use of a GFP-fusion genomic DNA library.";
RL Genes Cells 5:169-190(2000).
CC -I- FUNCTION: AFTER THE ONSET OF MITOSIS, AT MID- TO LATE ANAPHASE,
CC CO-LOCALIZES WITH THE MEDIAL ACTIN RING. MAY PLAY A ROLE IN
CC CYTOKINESIS.
CC -I- SUBCELLULAR LOCATION: Cytoplasmic.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC DR EMBL, AL133357; CAB62414.1; -.
CC DR EMBL, AF0130473; AAB65416.1; ALT_INIT.
CC DR EMBL, AB028012; BAA87316.1; -.
CC KW Coiled coil.
CC FT DOMAIN 57 361 COILED COIL (POTENTIAL).
CC FT DOMAIN 443 463 COILED COIL (POTENTIAL).
CC FT DOMAIN 542 740 COILED COIL (POTENTIAL).
CC FT DOMAIN 804 1106 COILED COIL (POTENTIAL).
CC FT DOMAIN 1223 1427 COILED COIL (POTENTIAL).
CC FT DOMAIN 1497 1555 COILED COIL (POTENTIAL).
CC FT DOMAIN 1601 1664 COILED COIL (POTENTIAL).
CC FT SEQUENCE 1727 AA; 197858 MW; F820BF8D9C132644 CRC64;
SO

Query Match 3.3%; Score 316; DB 1; Length 1727;
Best Local Similarity 20.1%; Pred. No. 5.2e-06;
Matches 369; Conservative 267; Mismatches 652; Indels 548; Gaps 84;

OY 7 RVEEVQVYNAMQKILECPICLEIKEPVSTKCDHIFCKFGLKLANO----- 54
DB 85 QAESRNLIISVLRN-----ENESLKTNLNENONKRFALPTENQSLRRANSELSQ 134
OY 55 -KKRPSQCPCKNDITKRSLO-ESTRFSQLYBELIKITCAQLODTGLEXYANSTYPAKKEN 112
DB 135 SKIASSEDLSTAKDDI--EALONENSHLGEVOYSAHQALSDIE---ERKKOHFA--- 183
OY 113 NSPEHLKDEVAIIISMGYRNRAKRLTQSEP-----ENPSLQETSLSVQSLNCTGYVT- 164
DB 184 SSSSRVKEFI-LVDEKSAVSDLASDHSKVCBEKLEVSSROYODELKKIAGLAQNT 242
OY 165 LRTKRIQIOPKTSYVIELGSDS---SEDPYNAKATYCSVGDQDELLQITPGTRDEISLDS 220
DB 243 LNEKIQIQLFEQKRNSVSDSGNISKILETDPTSIKEELEVEVEIQKRLTALMEKSSSELQSEV 302
OY 221 AKKAACESEFDPVNTNTEHHQPSNN--DLNTTEKRAAERHPEKYOGSSVNLHVEPGCNT 278
DB 303 AALOEKLTQSQSLYNNVTLEINNKKQOLLISENSLRELD- EKYD-SVYSELSQVYKRNKNT 360
OY 279 HASS-----LQENSSLLLTDRMNVKAEKCNK 307
DB 361 SVSAGVGLFSLAOKLSAVQNPERSFTKRVYSDNNKLOOKVYSLKLOLDRLTNKRSFCEQ 420
OY 308 SKQ--PGLARSQH-----NFWAGSKCTNDRRPSTETKQVLDNADPLCEKRE--- 352
DB 421 VKQKILPVYKQQRSELVNRNNTYMANFLSELETSSNNNTL-----KYQDELSTKMKQECYL 475
OY 353 -WNRQKLPCESENPRDTEDPWITLNTSSLOKYNWFNRS-----DELLGSDSDSHGE 402

```

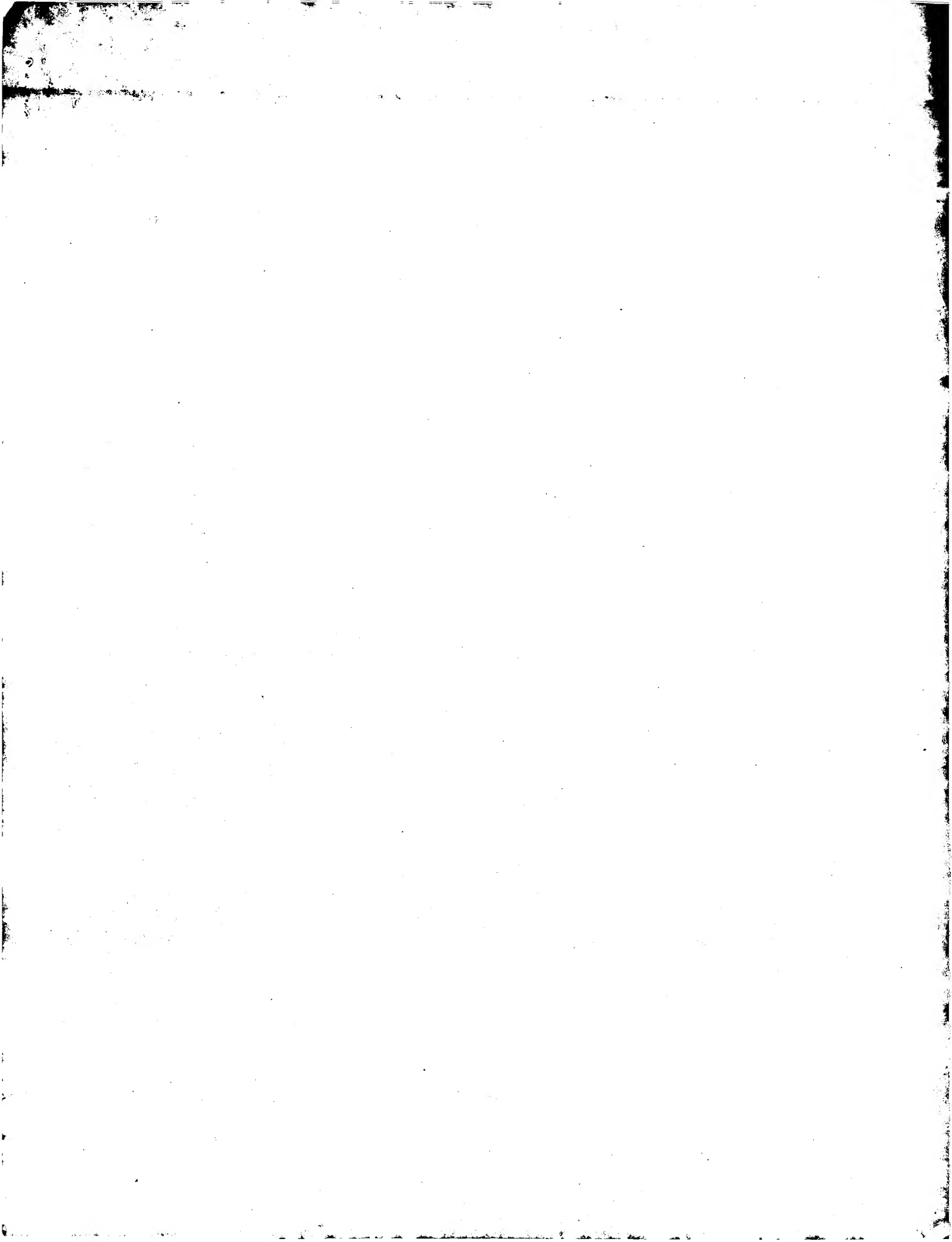
Db	476	QUTASRTQCS	DLSDREV-----ICLMAELDLHNETSRNVPATVOVALDE-YAQNSTASE	529
Qy	403	SESNAKVADVDL	VLNEVDYSGSSSEKIDLASDPHEALICKSERVHKSVSNEIDKIFG	462
Db	530	TLVNKELANFSS	IKAVSKTELEKRV-----RALECODE-10KOTVOQOISNAVKE	580
Qy	463	KYRRKAKSLPN	SIYHTENLIGAFTEQOIIORPLTKIKRRP-----TSLHRED	516
Db	581	NSNTLSEQIKNL	-----ESELNSKIKNSLNERLKLEMATSSSLSHNSXGNLDD	636
Qy	517	FIKKAD-----	LAVOKTPEMINGNOTGOTEONGVANNITNSGHEKTKG	559
Db	637	KKKSIDESTRE	LEKRYEYRNEMPAIQESLSKRNDL-LSEHEALRKELENSKYOOOLST	695
Qy	560	DSIQEKNENP	PIESLEKE-----SAFKTAKPEISSISNMELE	597
Db	696	DLRTVAN--	NDOVEAFKKEAKELRSTINOLDIISBQDORASKFEABELHVNSLABRLEGE	753
Qy	598	LNIHNSKAPKRR	LKRSKSTRIIHALE-LVYSRNLSPNCFILO--IDSCSSSEIKKKK	654
Db	754	LNA--SKGEKD-	LKKRTOERLISENDKLLAERERLMSLVSDLOFTLNOOOLSDAARYK	809
Qy	655	YNQMPVRISRN	LOIWEKGEPAATGAKSN-----KNE	686
Db	810	FES--EKESLS	SLCKLKE--SNEKMSNDLHSLQKSLKSGCIESSRKITLMEKQSLSE	865
Qy	687	QTSKRHDSDT	PBELKITNAPG-----SFTKCSNTS-----	716
Db	866	DNKRILDNOQ	MEITLOELNGVIELEKQFSTLEAKFTQOKNTYSERREALLESLDLO	925
Qy	717	---ELKEFVNP	SLPREK-----BEKLTGVVSNNAEDPKDMLSGERVLOTESV	764
Db	926	SKHTSLESQY	NSLERNIEQLQASKLAEMERYETVED-----EYRLDTSESL	974
Qy	765	ESSSILVPGT	YVGRQESISLEVSTLQAKTEPKVCVQCAAFENRPGILHGCSDKMRN	824
Db	975	ENHNLKIT--	---SLEQRLVILQ-----DEIASSSLRCGN-----ITKQSET	1011
Qy	825	DT---EGEGRY	PLGHEVHNSRETSTIEMESELDIAOYLONTFERVSK--ROSFALFSPNP	875
Db	1012	RVALLEENK	HL-LNNELSNR--NAEKQHEKENYKQOOLLVTFDCLKTREDY-----	1062
Qy	876	GNAEBECATF	SAHSGSLKKQSEPKVTECEQKEE-----NOGKNESIKRPQVQVNTAG	928
Db	1063	--EKELLRH	ADARSTLOKREDYKALEOVEDKLKEITALKAGINESO-----P	1108
Qy	929	FRVVOQKQK	PVNACSTIKGSRFL--SQOFGNETGLITPR--KHGL--LONP	977
Db	1109	FP-ISEKEDP	LQOEYVYLKQOAMLLTOLQSSNLFAFITSPDSDVSVMKLGLSDLONH	1167
Qy	978	Y-RIPPLP	IKSFVTKCKNNLEPFEDHSMPERENGNEIDSTVSTIR--NLIREN	1034
Db	1168	VKRISKEMLIS	-----CQOROLL--FILE-----NKKIKRTVASSNRVIADLORG	1209
Qy	1035	VFKRASSSN	INEVGST-----NEVGSISINEIGSSDENIAELGRNBPKLN--A	1082
Db	1210	ITERGVSTSE	VGERSNYLWVALNLNENSKSLRENLBRNEVITEL-REKIFETLKDTLA	1268
Qy	1083	MURGLVLOP	PEYKQK-----LPGSNKHPRIKKQETEEVYQVYNDPFSYLLSDNLEQ	1135
Db	1269	NFRLLKREDE	LOLETERKAAYKLENSN---EYKRNHOETILSINSSTST--	1315
Qy	1136	PMGSSHAYO	VOVCE--TPDDLLDD-----GEIK--EDTSFAENDIKESAAVFSKVOKE	1185
Db	1316	---SSDARL	NELNVLVSKNLELBELNOETIGHLSKSELEYVKSSEDLENRAONOSKIDOLE	1372
Qy	1186	LSRSPSTPT	HLAOGYR-----RGAKKLESEENMLSEDEDEL-----PCFO-HILF	1231
Db	1373	LK-----	NKRLAAMRTKYEOYVYKNSLEKHNQIDROLOSKTSELEKAVACHQOLBOL	1425
Qy	1232	GKVNNPISQ	STRH--STVATECLSKNTEENLLSKNSLDCSNOVYLAKASQEHNLSEET	1289

```

Db      1426 NKPSAIPYATTOSEPSTVSLIEFN-STKEELSTÖRKLEIND----- 1467
QY      1290 KCSASLFSSQCELEDLTANTNTQDPFLIGSSK-----QMRHOSQGVGLSDKEIV 1341
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1468 ----ILNTTKELEKVRQNSNKS-----GTSKDTFLPNEEMERKKVMQOEVLRLRSRI 1518
QY      1342 SDEBERGTGLENNQOEQSDSNLGEAAGCESETSVS----- 1379
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1519 AKELQKNEILLRKONQVLODOVKALQETVVSSEEAESASVHADTKDLENLKTEEMLSYTF 1578
QY      1380 ----EDCGSLSQSDILFT-----QORDTMOHN----- 1403
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1579 QVIFNESISDFSTADFTTFVOKEMEKREILQKDVQVQASHQKOLDNIRKLEMRN 1638
QY      1404 ---LIRLOQEMAELEAVLEQHGQSOPNSYPSIISDSSALEDLRNPQSTSEKAVLTSQ-- 1458
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1639 KKLKSLMEKLAIVRALEDSKKDS---PAIIS---LEASKNTDSNKSNSEVPAQYK 1691
QY      1459 ----KSSEYPISONPEGISADK-FEVSADSSTSK 1487
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1692 EKKLIAKTHSVDTNPPKRRSSPDAGMDVSDVKKAK 1727

```

Search completed: June 27, 2003, 10:23:22  
 Job time : 52 secs



	17	3452	35.8	944	6	09BDV2	09bdv2 lama glams
	18	3435	35.6	959	11	09ZAE3	09zae3 sciturus ni
	19	3419.5	35.4	963	11	099MX4	099mx4 glucomys
	20	3406.5	35.3	931	6	09BDM2	09bdm2 chaetophia
	21	3382	35.1	948	6	09BDU6	09bdu6 felis silve
	22	3361	34.8	942	6	09BDU8	09bdu8 physeler ca
	23	3356	34.8	940	6	09BDU7	09bdu7 megaptara n
	24	3345	34.7	759	4	09Z897	09z897 homo sapien
	25	3337.5	34.6	931	6	09BDU4	09bdu4 trichechus
	26	3336.5	34.6	949	6	09TT62	09tt62 hipposidero
	27	3318	34.4	934	6	08SOG2	08sg2 natalus stro
	28	3298	34.2	946	6	09BDV0	09bdv0 sus scrofa
	29	3294	34.1	938	6	09BDM0	09bdm0 bradypus tr
	30	3293	34.1	924	6	09BDU5	09bdas dugong dugo
	31	3283.5	34.0	945	6	09BDM1	09bdw1 lamandua te
	32	3276	34.0	936	6	09TT68	09tt68 myotis daub
	33	3263	33.8	942	6	09BDV5	09bdv5 manis sp. o
	34	3262.5	33.8	931	6	08MMT8	08mt8 sotolia flu
	35	3245	33.6	930	6	09BDU3	09bd3 loxodonta a
	36	3236	33.5	906	6	08SQF7	08sqf7 rhinopoma b
	37	3234	33.5	928	6	09BDU2	09bd2 elphas max
	38	3226	33.4	928	6	09BDU9	09bd9 hippopotamu
	39	3219	33.4	944	6	09BDV7	09bdv7 scalopus ag
	40	3217	33.3	930	6	09TT67	09tt67 tadarida br
	41	3197	33.1	928	6	09BDM3	09bdw3 dasypus nov
	42	3171.5	32.9	917	6	08MMT4	08mt4 rousettus a
	43	3171	32.9	935	6	09TT63	09tt63 pteropus ra
	44	3167	32.8	930	6	08SOG1	08sg1 rhinolophus
	45	3157.5	32.7	929	6	09BDT4	09bdt4 orycteropus

ALIGNMENTS

RESULT 1							
09GKK8	ID	09GKK8	PRELIMINARY:	PRT:	1863	AA:	
AC	09GKK8;	DT	01-MAR-2001 (TREMBlrel. 16, Created)				
DT	01-MAR-2001 (TREMBlrel. 16, last sequence update)						
DT	01-JUN-2002 (TREMBlrel. 21, last annotation update)						
DE	BRCAL (Fragment).						
CN	BRCAL.						
OS	Pan troglodytes (Chimpanzee).						
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Pan.						
OX	NCBI_TaxID=9598;						
RN	[1]						
RP	SEQUENCE FROM N.A.						
RA	TISSUE=BLOOD.						
RC	Takeda R., Hink R.L., Jogodka C., Walter N.A.R., Messier W.;						
RT	"Positive selection on the human BRCAL gene may have resulted from pressure for prolonged care for infants."						
RL	Submitted (NOV-1999) to the EMBL/GenBank/DBI databases.						
CC	-1 SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.						
DR	EMBL: AF207822; AAC43492.1; -						
DR	InterPro: IPR001357; BRCT.						
DR	InterPro: IPR002378; Brtl.cancerl.						
DR	InterPro: IPR001990; Granlin.						
DR	InterPro: IPR001841; znf_ring.						
DR	Pfam: PF00533; BRCT_2.						
DR	Pfam: PF00097; zf-C3HC4_1.						
DR	PRINTS: PRO0493; BRSTCANCER1.						
DR	SMART: SMO0292; BRCT_2.						
DR	SMART: SMO0184; RING_1.						
DR	-1 SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.						
DR	EMBL: AF207822; AAC43492.1; -						
DR	InterPro: IPR001357; BRCT.						
DR	InterPro: IPR002378; Brtl.cancerl.						
DR	InterPro: IPR001990; Granlin.						
DR	InterPro: IPR001841; znf_ring.						
DR	Pfam: PF00533; BRCT_2.						
DR	Pfam: PF00097; zf-C3HC4_1.						

FT NON\_TER 1863 1863  
SQ SEQUENCE 1863 AA; 207916 MW: A365EB5A34A77F4A CRC64;  
Query Match 98.2%; Score 9471; DB 6; Length 1863;  
Best Local Similarity 98.3%; Pred. No. 0;  
Matches 1832; Conservative 9; Mismatches 22; Indels 0; Gaps 0;

QY 1 MDLSALRVEYONVINAMOKILECPICLLEKEPVSTKCDHIFCKCKMLKLNQKKGPQ 60  
DB 1 MDLSALRVEYONVINAMOKILECPICLLEKEPVSTKCDHIFCKCKMLKLNQKKGPQ 60  
QY 61 CPLCNDITKRSLOSTRPSQLYEBELKICAFOLDTGLEYANSTYFAKKENNSPRLKD 120  
DB 61 CPLCNDITKRSLOSTRPSQLYEBELKICAFOLDTGLEYANSTYFAKKENNSPRLKD 120  
QY 121 EYSIIOSMGYRRARARLLOSEPENPSLOETSLSVOLSNLGTVRTLTOKRIOPQKSVYI 180  
DB 121 EYSIIOSMGYRRARARLLOSEPENPSLOETSLSVOLSNLGTVRTLTOKRIOPQKSVYI 180  
QY 181 ELGSDSEDTYVKATYCSVGDELQITPQGTREISLDSAKKAACEFSETDVTNTHIQ 240  
DB 181 ELGSDSEDTYVKATYCSVGDELQITPQGTREISLDSAKKAACEFSETDVTNTHIQ 240  
QY 241 PSNNDLNTTEKRAARHPEKYOGSSVSNLHVPCGNTTHASSLOHENSLLTKDMNVE 300  
DB 241 PSNNDLNTTEKRAARHPEKYOGSSVSNLHVPCGNTTHASSLOHENSLLTKDMNVE 300  
QY 301 KAEFCNKSKOPCLARSOHNRWAGSKETCDRRTPTSEKTVDLNADPLCRKEKNKOKLPC 360  
DB 301 KAEFCNKSKOPCLARSOHNRWAGSKETCDRRTPTSEKTVDLNADPLCRKEKNKOKLPC 360  
QY 361 SENPRDTEVPWITLNSIOKVNEWFSRSDDELGSDSDSHGSESNKAVADYLDVLEND 420  
DB 361 SENPRDTEVPWITLNSIOKVNEWFSRSDDELGSDSDSHGSESNKAVADYLDVLEND 420  
QY 421 EYSGSSEKIDLLASPHHEALICKSERVHSKSVESNEDKIFGKYTRKRSASLPLSHVTN 480  
DB 421 EYSGSSEKIDLLASPHHEALICKSERVHSKSVESNEDKIFGKYTRKRSASLPLSHVTN 480  
QY 481 LIIGAFVEPOLIOBRPLTNKLRKRRTPTSGLHPEDEIKKADLAOKTPEMINOGNOTE 540  
DB 481 LIIGAFVEPOLIOBRPLTNKLRKRRTPTSGLHPEDEIKKADLAOKTPEMINOGNOTE 540  
QY 541 ONGQVMTNTSGHENKTGDSIONENKPNPIESLEESAFTKAEPDISSISNMELENI 600  
DB 541 ONGQVMTNTSGHENKTGDSIONENKPNPIESLEESAFTKAEPDISSISNMELENI 600  
QY 601 HNSKAPKKNRLRKSTRIHALVELVSNLSPPNCTELOIDSCSSSEIKKKKNQMY 660  
DB 601 HNSKAPKKNRLRKSTRIHALVELVSNLSPPNCTELOIDSCSSSEIKKKKNQMY 660  
QY 661 RHRSLQLOMEGKEPATKSNKPNQOTSKRHSDTPPELKLINAAGSTKCSNTELKE 720  
DB 661 RHRSLQLOMEGKEPATKSNKPNQOTSKRHSDTPPELKLINAAGSTKCSNTELKE 720  
QY 721 FVNPSLPREEKEKLETYKVSNNADPKDMLSGERVLOTERTSVSSSISLPGTDYGTQ 780  
DB 721 FVNPSLPREEKEKLETYKVSNNADPKDMLSGERVLOTERTSVSSSISLPGTDYGTQ 780  
QY 781 EYSISLEVSTIGKAKTEPNKCYSCAEPENKGLIHGCKDNNDTEGKYYLGHVNH 840  
DB 781 EYSISLEVSTIGKAKTEPNKCYSCAEPENKGLIHGCKDNNDTEGKYYLGHVNH 840  
QY 841 RETSIEMESESLDAYLONTFVSKROSFALFNSGNAEBCATSAHSGSLKKOSPKYT 900  
DB 841 RETSIEMESESLDAYLONTFVSKROSFALFNSGNAEBCATSAHSGSLKKOSPKYT 900  
QY 901 FECDKEENQCKNESNIKPVQVTNTAGFPVVGOKDPVDNAKCSIKGSRFCLSSQFNG 960  
DB 901 FECDKEENQCKNESNIKPVQVTNTAGFPVVGOKDPVDNAKCSIKGSRFCLSSQFNG 960  
QY 961 NETGLITPNKGLLONPYRIPPLPIKSFVTKCKNLLENFEHSHMSPEREMNGENIT 1020  
DB 961 NETGLITPNKGLLONPYRIPPLPIKSFVTKCKNLLENFEHSHMSPEREMNGENIT 1020

DB 961 NETGLITPNKGLLONPYRIPPLPIKSFVTKCKNLLENFEHSHMSPEREMNGENIT 1020  
QY 1021 STVSTISRRNNIRENVFKKASSNINEVGSSSTNEVGSSINIEGSSPDENIOALGRRNGPKL 1080  
DB 1021 STVSTISRRNNIRENVFKKASSNINEVGSSSTNEVGSSINIEGSSPDENIOALGRRNGPKL 1080  
QY 1081 NAMRLGLVLOPEVYKOSLPGSNCKHPEIKKOEEVQVTNTDFSPYLLISDNLEQPMGSS 1140  
DB 1081 NAMRLGLVLOPEVYKOSLPGSNCKHPEIKKOEEVQVTNTDFSPYLLISDNLEQPMGSS 1140  
QY 1141 HASQVSETPDLDLDGELKEDTSAENDIKESSAVFSKSVQKGLSPPSPFTHTHLAQ 1200  
DB 1141 HASQVSETPDLDLDGELKEDTSAENDIKESSAVFSKSVQKGLSPPSPFTHTHLAQ 1200  
QY 1201 GYRGAKKLESSEENISDEELPCFOHLIFKVNINISQSTRHSTVATECLSKNTEENL 1260  
DB 1201 GYRGAKKLESSEENISDEELPCFOHLIFKVNINISQSTRHSTVATECLSKNTEENL 1260  
QY 1261 LSLKNSLNDCSNOVILAKASQEHLSSEETKCSASLFSQCSLEDLTANTNTQDPFLIGS 1320  
DB 1261 LSLKNSLNDCSNOVILAKASQEHLSSEETKCSASLFSQCSLEDLTANTNTQDPFLIGS 1320  
QY 1321 SKQMRHQSOGVGLSDKELVSDDEERGTLSENNQEQSDMSNNGEASAQCESETSYSE 1380  
DB 1321 SKQMRHQSOGVGLSDKELVSDDEERGTLSENNQEQSDMSNNGEASAQCESETSYSE 1380  
QY 1381 DCSGLSSQSDILTTQOORIMQNLIRKLOEMAEFLAVLEOHGSOQSPNSPTSIDSSALE 1440  
DB 1381 DCSGLSSQSDILTTQOORIMQNLIRKLOEMAEFLAVLEOHGSOQSPNSPTSIDSSALE 1440  
QY 1441 DLNRPQOSTSEKAVLTSOKSSEYPTSONPEGLSADKFEVADSSTSKNKEGVERSSPSK 1500  
DB 1441 DLNRPQOSTSEKAVLTSOKSSEYPTSONPEGLSADKFEVADSSTSKNKEGVERSSPSK 1500  
QY 1501 CPSLDDRWYHSCSGLQNRNYPQSOEELIKYVDVEEQLESGBPDLETYSYLRDQLEG 1560  
DB 1501 CPSLDDRWYHSCSGLQNRNYPQSOEELIKYVDVEEQLESGBPDLETYSYLRDQLEG 1560  
QY 1561 TYLESGLISLFSDDPESDEPESARAVGNIPSTSAKVYPOLKVAESQSPAHAHTT 1620  
DB 1561 TYLESGLISLFSDDPESDEPESARAVGNIPSTSAKVYPOLKVAESQSPAHAHTT 1620  
QY 1621 DTAGYNAMESVSREKPELTASTERVNKRMSNVASGLTPEEFMLVYKFAKRNHITLNL 1680  
DB 1621 DTAGYNAMESVSREKPELTASTERVNKRMSNVASGLTPEEFMLVYKFAKRNHITLNL 1680  
QY 1681 TEETHYVAKTDAEFVCERTLYFLGIAQKVVSYFWVTOSIKERKMLNHDDEVRGDV 1740  
DB 1681 TEETHYVAKTDAEFVCERTLYFLGIAQKVVSYFWVTOSIKERKMLNHDDEVRGDV 1740  
QY 1741 VNGRNHOGPRARRESQDRKIFRGLICCYGPTNNPTQOLBMMVQOLGASVYKELSSFTL 1800  
DB 1741 VNGRNHOGPRARRESQDRKIFRGLICCYGPTNNPTQOLBMMVQOLGASVYKELSSFTL 1800  
QY 1801 GTGVHPHVVQPDANTEDNGFHAIGQMCBAVYTRREVLDSVALYQCOELDTYLIPIPIH 1860  
DB 1801 GTGVHPHVVQPDANTEDNGFHAIGQMCBAVYTRREVLDSVALYQCOELDTYLIPIPIH 1860  
QY 1861 SHY 1863  
DB 1861 SHY 1863

RESULT 2  
ID 015129 PRELIMINARY; PRT: 1792 AA.  
AC 015129;  
DT 01-JAN-1998 (Tremblrel. 05, Created)  
DT 01-JAN-1998 (Tremblrel. 05, last sequence update)  
DE 01-JUN-2001 (Tremblrel. 17, last annotation update)  
DE Breast and ovarian cancer susceptibility protein splice variant.  
GN BRCA1.  
OS Homo sapiens (Human).



OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OX Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NC NCBITaxID:9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=TESTIS;  
 RX MEDLINE=96172829; PubMed=8569721;  
 RA Holt J.T., Thompson M.E., Szabo C., Robinson-Benton C., Arteaga C.L.,  
 RA King M.C., Jensen R.A.;  
 RT "Growth retardation and tumour inhibition by BRCA1.";  
 RL Nat. Genet. 12:298-302(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=TESTIS;  
 RA Holt J.T., Robinson-Benton C.;  
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF005068; AAB61673.1; -;  
 DR InterPro: IPR001357; BRCT.  
 DR InterPro: IPR002378; Brst\_cancer1.  
 DR InterPro: IPR001990; Granin.  
 DR Pfam: PF00533; BRCT; 2.  
 DR PRINTS: PR00493; BRSTCANCER1.  
 DR SMART: SM00292; BRCT; 2.  
 DR PROSITE: PS00172; BRCT; 2.  
 DR PROSITE: PS00422; GRANINS\_1; UNKNOWN\_1.  
 SQ SEQUENCE 1792 AA; 199692 MW; 0CB9B06C0791A4A4 CRC64;

Query Match 95.7%; Score 9233.5; DB 4; Length 1792;  
 Best Local Similarity 99.8%; Pred. No. 0;

Matches 1788; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 72 SIQESTRFQSLVEELLKICAFOLDGLEYANSYNFAKKENNSPEILKDEVTIISGMGR 131  
 DB 2 SIQESTRFQSLVEELLKICAFOLDGLEYANSYNFAKKENNSPEILKDEVTIISGMGR 61  
 QY 132 NNAKRLIOSEPNPSLQETSLSVQSLNIGTVTLTKRIQPKTSVYIELGSDSEDTV 191  
 DB 62 NNAKRLIOSEPNPSL-ETSLSVQSLNIGTVTLTKRIQPKTSVYIELGSDSEDTV 120  
 QY 192 NKATYCSVGDDELOTTPOGTTRDEISLSAKKAACEFSETDVTNHHOQPSNNDLTTTK 251  
 DB 121 NKATYCSVGDDELOTTPOGTTRDEISLSAKKAACEFSETDVTNHHOQPSNNDLTTTK 180  
 QY 252 RAERHPEKYQGSVSNLHVEPCGTNTAASLQIHENSLLLTKDRNNVKAKEFCNKSOP 311  
 DB 181 RAERHPEKYQGSVSNLHVEPCGTNTAASLQIHENSLLLTKDRNNVKAKEFCNKSOP 240  
 QY 312 GLARSHNHWAGSKETCNDRRTPTSTKAYDLNADPLCERKENNKKLPCCSENPROTDEVP 371  
 DB 241 GLARSHNHWAGSKETCNDRRTPTSTKAYDLNADPLCERKENNKKLPCCSENPROTDEVP 300  
 QY 372 WITLNSIOKVNMWFSRDELLGSDDSHDGSESNKAVADVLDVNEVDEYSGSSSEKIDL 431  
 DB 301 WITLNSIOKVNMWFSRDELLGSDDSHDGSESNKAVADVLDVNEVDEYSGSSSEKIDL 360  
 QY 432 LASDPHEALICKSERVHKSVESENIEDKIFGKTYRKKASLPMLSHVTENLIIIGAFTEPO 491  
 DB 361 LASDPHEALICKSERVHKSVESENIEDKIFGKTYRKKASLPMLSHVTENLIIIGAFTEPO 420  
 QY 492 TIQERPLTNLKRKRRTPTSGLHPEDFIKKADLAVALQTPMINOQNTQEQVNNITNS 551  
 DB 421 TIQERPLTNLKRKRRTPTSGLHPEDFIKKADLAVALQTPMINOQNTQEQVNNITNS 480  
 QY 552 GHENKTGGSIONENPNPISLEKESAKTKAEPISISSISNMELNLNHNKAPKKNL 611  
 DB 481 GHENKTGGSIONENPNPISLEKESAKTKAEPISISSISNMELNLNHNKAPKKNL 540  
 QY 612 RRRKSTRHIALELVYSRNLSPNCTELOIDSCSSSEELIKKKKYNQMPRHRSNRLQMEG 671  
 DB 541 RRRKSTRHIALELVYSRNLSPNCTELOIDSCSSSEELIKKKKYNQMPRHRSNRLQMEG 600  
 QY 672 KEPATGAKSKNPNEDTSKRHSDTFPELKLINAQSGFTKCSNTSELKEFVNSLPREEK 731  
 DB 731 KEPATGAKSKNPNEDTSKRHSDTFPELKLINAQSGFTKCSNTSELKEFVNSLPREEK 731

DB 601 KEPATGAKSKNPNEDTSKRHSDTFPELKLINAQSGFTKCSNTSELKEFVNSLPREEK 660  
 QY 732 EEKLETVKSNNAEDKDLMLSGERYLOTERSVSSSISLVPTDGTQESTISLLEVSRL 791  
 DB 661 EEKLETVKSNNAEDKDLMLSGERYLOTERSVSSSISLVPTDGTQESTISLLEVSRL 720  
 QY 792 GAKTEPNKCVSOGAFAFNPBKGLIHGCSKDNNDTGFYKPYPLGHEVNHSHETSIEMEESE 851  
 DB 721 GAKTEPNKCVSOGAFAFNPBKGLIHGCSKDNNDTGFYKPYPLGHEVNHSHETSIEMEESE 780  
 QY 852 LDAQYLONTFKVSKROSFAFNPBGNAEECATFSAHSGSLKKQSPKVFTECQKEBNQ 911  
 DB 781 LDAQYLONTFKVSKROSFAFNPBGNAEECATFSAHSGSLKKQSPKVFTECQKEBNQ 840  
 QY 912 KNSNNTKPVQTVNTTGFPPVGGOKDPVDNAKSLIGSGSFCLSSQFRGETLTPNKH 971  
 DB 841 KNSNNTKPVQTVNTTGFPPVGGOKDPVDNAKSLIGSGSFCLSSQFRGETLTPNKH 900  
 QY 972 GLQNPYRIPPLPIKSFYKTKCKNMLENPEHSHMSPERENGNEIPSTVSTISRNNI 1031  
 DB 901 GLQNPYRIPPLPIKSFYKTKCKNMLENPEHSHMSPERENGNEIPSTVSTISRNNI 960  
 QY 1032 RENWFEKASSNINEVGSSTNEVGSSINETIGSDENIQALGRNRPKLNAMLRLGLVLP 1091  
 DB 961 RENWFEKASSNINEVGSSTNEVGSSINETIGSDENIQALGRNRPKLNAMLRLGLVLP 1020  
 QY 1092 EYVKOSLPGSNCKHPRIKKOEYEVQVNTDPSPLIDNLEQPMGSSHASVGCETPD 1151  
 DB 1021 EYVKOSLPGSNCKHPRIKKOEYEVQVNTDPSPLIDNLEQPMGSSHASVGCETPD 1080  
 QY 1152 DILLDGEIKEDTFFAANDIKESSAVFSKSVOKELRSRSPFHTHIAQYRGAKKLES 1211  
 DB 1081 DILLDGEIKEDTFFAANDIKESSAVFSKSVOKELRSRSPFHTHIAQYRGAKKLES 1140  
 QY 1212 SEENLSEDEBELPCFOHLLFGKVNNTPSQSTRHSTVATECLSKNTEENLILKNSINDCS 1271  
 DB 1141 SEENLSEDEBELPCFOHLLFGKVNNTPSQSTRHSTVATECLSKNTEENLILKNSINDCS 1200  
 QY 1272 NOVYILAKASOEHNLSETKCSASFSSQSELEDLANTQOPLIGSSKOMRHOSEGO 1331  
 DB 1201 NOVYILAKASOEHNLSETKCSASFSSQSELEDLANTQOPLIGSSKOMRHOSEGO 1260  
 QY 1332 GVLGSDKELVSDDEEGTGLLENNQEQSMDNLGEAASGCESTVSEDCSGLSSQSDI 1391  
 DB 1261 GVLGSDKELVSDDEEGTGLLENNQEQSMDNLGEAASGCESTVSEDCSGLSSQSDI 1320  
 QY 1392 LTTQORDTQOHNLIKIQOEMALEAVLEQHGSPNSYPSIISDSSALEDLRNPPOSTSE 1451  
 DB 1321 LTTQORDTQOHNLIKIQOEMALEAVLEQHGSPNSYPSIISDSSALEDLRNPPOSTSE 1380  
 QY 1452 KAVLTQSKSEXPISQNPGLSADKFEVSADSTSNKKEGVRSSPKCPSLDDRWYMH 1511  
 DB 1381 KAVLTQSKSEXPISQNPGLSADKFEVSADSTSNKKEGVRSSPKCPSLDDRWYMH 1440  
 QY 1512 SCGSLQNNRNPQOEBLIVVDVEEQOLESQPHDLETSTYLRQDLEGTPLYLSCISLF 1571  
 DB 1441 SCGSLQNNRNPQOEBLIVVDVEEQOLESQPHDLETSTYLRQDLEGTPLYLSCISLF 1500  
 QY 1572 SDDPESDPEEDRAPESARVGNTPSSTALKVPOLKVAESASQSPAAAHHTDTAGYNAMEES 1631  
 DB 1501 SDDPESDPEEDRAPESARVGNTPSSTALKVPOLKVAESASQSPAAAHHTDTAGYNAMEES 1560  
 QY 1632 VSRKPELTASTERVKKRMSVVSGLTPPEPMVYVFAKHHHTLNLNLTTEETHVYMT 1691  
 DB 1561 VSRKPELTASTERVKKRMSVVSGLTPPEPMVYVFAKHHHTLNLNLTTEETHVYMT 1620  
 QY 1692 DAEFVCEKTLKFLGIAGKVVVSFYFWYQSIKERKMLNEHDFEYGVADVNGNHHGPR 1751  
 DB 1621 DAEFVCEKTLKFLGIAGKVVVSFYFWYQSIKERKMLNEHDFEYGVADVNGNHHGPR 1680  
 QY 1752 ARESDOKTFRGLETICCPETNMPDQLEMMVQLCGASVYKELSFLLGTGVHPITVVO 1811  
 DB 1681 ARESDOKTFRGLETICCPETNMPDQLEMMVQLCGASVYKELSFLLGTGVHPITVVO 1740

QY 1812 PDANTDNGFHAIGMCEAPVYTRREVLDSVALYCCOELDTYLLIPOIPHSY 1863  
 |||||||  
 Db 1741 PDANTDNGFHAIGMCEAPVYTRREVLDSVALYCCOELDTYLLIPOIPHSY 1792

## RESULT 3

046485 PRELIMINARY: PRT: 1141 AA.  
 ID 046485  
 AC 046485:  
 DT 01-JUN-1998 (TReMBLrel. 06, Created)  
 DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)  
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)  
 DE Breast cancer type 1 susceptibility protein (Breast and ovarian cancer susceptibility protein) (Fragment).  
 GN BRCAL.  
 OS Gorilla gorilla gorilla (lowland gorilla).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Gorilla.  
 OX NCBI\_TaxID=9595;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98122577; PubMed=9462745;  
 RA Hacia J.G., Makalowski W., Edgemon K., Erdos M.R., Robbins C.M.,  
 RA Fodor S.P.A., Brody L.C., Collins F.S.;  
 RT "Evolutionary sequence comparisons using high-density oligonucleotide  
 arrays";  
 RL Nat. Genet. 18:155-158(1998).  
 CC -1- FUNCTION: NOT KNOWN. MAY REGULATE GENE EXPRESSION. MAY PLAY A ROLE  
 IN CELL PROLIFERATION AND DIFFERENTIATION (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).  
 DR EMBL: AF019076; AAC39584.1;  
 DR InterPro: IPR002378; Brst\_cancerl.  
 DR PRINTS: PR00493; BRSTCANCERL.  
 KM DNA-binding; Nuclear protein; Anti-oncogene.  
 FT NON\_TER 1  
 FT DOMAIN 277 284 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
 FT DOMAIN 427 430 POLY-LYS.  
 FT DOMAIN 985 1006 LEUCINE-ZIPPER (POTENTIAL).  
 FT NON\_TER 1141 1141  
 SQ SEQUENCE 1141 AA; 127339 MW; A09531803BB96C3D CRC64;

Query Match 60.1%; Score 5795; DB 6; Length 1141;  
 Best Local Similarity 98.1%; Pred. No. 3,8e-291;  
 Matches 1119; Conservative 6; Mismatches 16; Indels 0; Gaps 0;

QY 225 ACESEPDVNTTEHNPNSNLTNTEKRAAEHREPKYGGSSVSNLHVEPCGNTHTASSIQ 284  
 |||||||  
 Db 1 ACESERDVNTTEHNPNSNLTNTEKRAAEHREPKYGGSSVSNLHVEPCGNTHTASSIQ 60  
 QY 285 HENSSLLTKDRMVNEKAFCNKSQKPGIARSQHNRNAGSKETCDNRTPTEKKVLDNA 344  
 |||||||  
 Db 61 HENSSLLTKDRMVNEKAFCNKNKQGLARSQHNRNAGSKETCDNRTPTEKKVLDNA 120  
 QY 345 DPLCEKEMKQKLPCESENPRTEDVPWITLNSIQKYNEMFSRDELLGSDSDGSE 404  
 |||||||  
 Db 121 DPLCEKEMKQKLPCESENPRTEDVPWITLNSIQKYNEMFSRDELLGSDSDGSE 180  
 QY 405 SNAKAVADVLTNVEDEYSGSSEKIDLLASDPHEALIKKSEVHKSVESENEDKIFGKT 464  
 |||||||  
 Db 181 SNAKAVADVLTNVEDEYSGSSEKIDLLASDPHEALIKKSEVHKSVESENEDKIFGKT 240  
 QY 465 YRRKASLPNLSHVTENLIGAFVTEPQIIQERPLTNLKRKRRTSGLHPEDFIKADLA 524  
 |||||||  
 Db 241 YRRKASLPNLSHVTENLIGAFVTEPQIIQERPLTNLKRKRRTSGLHPEDFIKADLA 300  
 QY 525 VQKPEMINOCTNOTEGNGVYMTNSGHEKTKGDSIQNEKNPPIISLEKESAFKTKA 584  
 |||||||  
 Db 301 VQKPEMINOCTNOTEGNGVYMTNSGHEKTKGDSIQNEKNPPIISLEKESAFKTKA 360  
 QY 585 EPISSISNMELTNHNSKAPKKNRKRKSTRHIALELVSNLSPPNCTEIQDSC 644  
 |||||||  
 Db 361 EPISSISNMELTNHNSKAPKKNRKRKSTRHIALELVSNLSPPNCTEIQDSC 420

QY 645 SSSSEIKKKKYNQMPVRHSRLQLMGKEPATGAKSKNPNEDQTSKRHSDTPPELKTN 704  
 |||||||  
 Db 421 SSSSEIKKKKYNQMPVRHSRLQLMGKEPATGAKSKNPNEDQTSKRHSDTPPELKTN 480

QY 705 APGFTKCSNTSELKEFVNPSCI.PREEKEKELETYKVSNNADPPDLMLSGERVQTERSV 764  
 |||||||  
 Db 481 APGFTKCSNTSELKEFVNPSCI.PREEKEKELETYKVSNNADPPDLMLSGERVQTERSV 540

QY 765 BSSSISLVPGTDYQGESISLEVESTLGKATPEPKVCVSOCAEENPGGLHGCSSKNRN 824  
 |||||||  
 Db 541 BSSSISLVPGTDYQGESISLEVESTLGKATPEPKVCVSOCAEENPGGLHGCSSKNRN 600

QY 825 DTEGFRYPDLGHEVNSRSTSIEMESSELDAYLQNTFKVSKROSFALFNSGNAEECAT 884  
 |||||||  
 Db 601 DTEGFRYPDLGHEVNSRSTSIEMESSELDAYLQNTFKVSKROSFALFNSGNAEECAT 660

QY 885 FSAHSGSLKKQSPKVTFCQKEENQGNESNIKPVQTVNTAGPPVYGQDKRPVDNAKC 944  
 |||||||  
 Db 661 FSAHSGSLKKQSPKVTFCQKEENQGNESNIKPVQTVNTAGPPVYGQDKRPVDNAKC 720

QY 945 STKGSRCTLSQFPGNETGLTPPKHGLQNPRIPLPFIKSPVTKCKNLLENFE 1004  
 |||||||  
 Db 721 STKGSRCTLSQFPGNETGLTPPKHGLQNPRIPLPFIKSPVTKCKNLLENFE 780

QY 1005 EHSMSPEREMGNENIPSTVSTISRNNIRENVFEKASSNINEVGSSTNEVGSSINEIGSS 1064  
 |||||||  
 Db 781 EHSMSPEREMGNENIPSTVSTISRNNIRENVFEKASSNINEVGSSTNEVGSSINEIGSS 840

QY 1065 DENIOAELGRNRPGLNMLRLGLVQPEVYQSLDPSGCKIPETIKQBEVQTVNTDF 1124  
 |||||||  
 Db 841 DENIOAELGRNRPGLNMLRLGLVQPEVYQSLDPSGCKIPETIKQBEVQTVNTDF 900

QY 1125 SPYLTSDNLEQPMGSSHASQVCSSTPDLDDGELKEPTSAENDIKESSAVFSQVQK 1184  
 |||||||  
 Db 901 SPYLTSDNLEQPMGSSHASQVCSSTPDLDDGELKEPTSAENDIKESSAVFSQVQK 960

QY 1185 ELRSRSPFTHTHLAGYRRGAKKLESSEENLSSEDEELPCFOHLLFGKVNIPISOSTRH 1244  
 |||||||  
 Db 961 ELRSRSPFTHTHLAGYRRGAKKLESSEENLSSEDEELPCFOHLLFGKVNIPISOSTRH 1020

QY 1245 STVATECISKNTENELSLKNSLNDCSNQVTLAAQSEHHLSEETKCSASLFSQGELE 1304  
 |||||||  
 Db 1021 STVATECISKNTENELSLKNSLNDCSNQVTLAAQSEHHLSEETKCSASLFSQGELE 1080

QY 1305 DLTANTNODPFLIGSSKQMRHOSQGVGLSDKEIVSDDERGTGLENNQEQSDSN 1364  
 |||||||  
 Db 1081 DLTANTNODPFLIGSSKQMRHOSQGVGLSDKEIVSDDERGTGLENNQEQSDSN 1140

QY 1365 L 1365  
 ||  
 Db 1141 L 1141

RESULT 4  
 046484 PRELIMINARY: PRT: 1141 AA.  
 ID 046484  
 AC 046484:  
 DT 01-JUN-1998 (TReMBLrel. 06, Created)  
 DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)  
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)  
 DE Breast cancer type 1 susceptibility protein (Breast and ovarian cancer susceptibility protein) (Fragment).  
 GN BRCAL.  
 OS Pan troglodytes (Chimpanzee).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Pan.  
 OX NCBI\_TaxID=9598;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98122577; PubMed=9462745;  
 RA Hacia J.G., Makalowski W., Edgemon K., Erdos M.R., Robbins C.M.,  
 RA Fodor S.P.A., Brody L.C., Collins F.S.;

```

RT      "Evolutionary sequence comparisons using high-density oligonucleotide
RT      arrays."
RL      Nat. Genet. 18:155-158(1998).
CC      -1- FUNCTION: NOT KNOWN, MAY REGULATE GENE EXPRESSION. MAY PLAY A ROLE
CC      IN CELL PROLIFERATION AND DIFFERENTIATION (BY SIMILARITY).
CC      -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC      EMBL: AF019075; AAC39583.1; -.
CC      InterPro: IPR002378; Brst_cancer1.
DR      PRINTS: PR00493; BRSTCANCER1.
KM      DNA-binding: Nuclear protein; Anti-oncogene.
FT      NON_TER 1
FT      DOMAIN 277 284 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT      DOMAIN 427 430 POLY-LYS.
FT      DOMAIN 505 509 POLY-GLU.
FT      DOMAIN 985 1006 LEUCINE-ZIPPER (POTENTIAL).
FT      NON_TER 1141 1141
SQ      SEQUENCE 1141 AA; 127399 MW; D72660B3CA632D5 CRC64;

Query Match      59.9%; Score 5776; DB 6; Length 1141;
Best Local Similarity 97.9%; Pred. No. 3,6e-290;
Matches 117; Conservative 6; Mismatches 18; Indels 0; Gaps 0;

OY      225 ACDEFSTDTNTNHHQPSNNDLNTTEKRAERHPEKYQSSVSNLHVEPCGNTTHASSIQ 284
DB      1 ACEFSETDVTNTEHHQPSNNDLNTTEKRAERHPEKYQSSVSNLHVEPCGNTTHASSIQ 60

OY      285 HENSLLTLTKDMNNVKAFCNKSQOPGLARSOHNWAGSKECNDRRPTSEKKVDLNA 344
DB      61 HENSLLTLTKDMNNVKAFCNKSQOPGLARSOHNWAGSKECNDRRPTSEKKVDLNA 120

OY      345 DPLCEKERNKOKLPCESENPRDTEVPWITLNSIQKVMFESRSDSLGSDSDHGESE 404
DB      121 DPLCEKERNKOKLPCESENPRDTEVPWITLNSIQKVMFESRSDSLGSDSDHGESE 180

OY      405 SNAKADVDVLYNEVDYSGSSEKIDLASDPHEALICSEKRVHKSQSVSNIEDKIFGKT 464
DB      181 SNAKADVDVLYNEVDYSGSSEKIDLASDPHEALICSEKRVHKSQSVSNIEDKIFGKT 240

OY      465 YRKKASLPVLSHVTENLIGAFVTEPQIOERPLNKLKRKRPTSGLPEDFIKKADIA 524
DB      241 YRKKASLPVLSHVTENLIGAFVTEPQIOERPLNKLKRKRPTSGLPEDFIKKADIA 300

OY      525 VQKTPMINQNGTQEQNGQVNNITNSGHNKTKGDSIONENPNPIESLEKESAFKTKA 584
DB      301 VQKTPMINQNGTQEQNGQVNNITNSGHNKTKGDSIONENPNPIESLEKESAFKTKA 360

OY      585 EPISSISIMELNLTINHSKAPKKNLRKSSRTHIALELVYSRNLSPNCTEIQIDSC 644
DB      361 EPISSISIMELNLTINHSKAPKKNLRKSSRTHIALELVYSRNLSPNCTEIQIDSC 420

OY      645 SSSEELKKKKYQMPVRRSRNLQIMBGKRPATGAKSKSNKPNQTSKRHSDTFPELKLTN 704
DB      421 SSSEELKKKKYQMPVRRSRNLQIMBGKRPATGAKSKSNKPNQTSKRHSDTFPELKLTN 480

OY      705 AGSFETKCSNTSELKEFNPSLPREEKEKLETVKSNNAEPKDKMLGSEKVLQTERSV 764
DB      481 AGSFETKCSNTSELKEFNPSLPREEKEKLETVKSNNAEPKDKMLGSEKVLQTERSV 540

OY      765 ESSSISLVPGTDYGTQOESISLLEVSTLGAKTEPNKCVSOCAAFENPKGLIGCSKDN 824
DB      541 ESSSISLVPGTDYGTQOESISLLEVSTLGAKTEPNKCVSOCAAFENPKGLIGCSKDN 600

OY      825 DREGFKYPLGHEVNSRSTISMESESLDAOYLQNTFKSKOSFALFSPNNAEBCCT 884
DB      601 DREGFKYPLGHEVNSRSTISMESESLDAOYLQNTFKSKOSFALFSPNNAEBCCT 660

OY      885 FSAHSGSLKQSPKVTFFCEQKLENGKKNESIKPQVNTIAGFPVVOXKRPVDNAKC 944
DB      661 FSAHSGSLKQSPKVTFFCEQKLENGKKNESIKPQVNTIAGFPVVOXKRPVDNAKC 720

OY      945 SIKGSRFLSSQFRGNETGLITTPNKHGLQNPYRIPPLPIKSEVYKTKCKKNLLENPE 1004
DB      721 SIKGSRFLSSQFRGNETGLITTPNKHGLQNPYRIPPLPIKSEVYKTKCKKNLLENPE 780

```

```

OY      1005 EHSMSPERENGNIPISTVSTISRNNIRENVFEKASSNINEVGSSITNEVGSS 1064
DB      781 EHSMSPERENGNIPISTVSTISRNNIRENVFEKASSNINEVGSSITNEVGSS 840

OY      1065 DENIQAEILGRNRPKLNMLRLGLVQPEVYKQSLPSSNCKHPEIKQEEVYQVNTDF 1124
DB      841 DENIQAEILGRNRPKLNMLRLGLVQPEVYKQSLPSSNCKHPEIKQEEVYQVNTDF 900

OY      1125 SPLIISDNLQPPGSSHASQVSEYTDLDLQGEIKEDTSPFANDIKESSAVSKSVQK 1184
DB      901 SPLIISDNLQPPGSSHASQVSEYTDLDLQGEIKEDTSPFANDIKESSAVSKSVQK 960

OY      1185 ELSRSPPTHTHLAGYRGAKKLESSEENLSEDEELPCROHLFGKVNIPQOSTTH 1244
DB      961 ELSRSPPTHTHLAGYRGAKKLESSEENLSEDEELPCROHLFGKVNIPQOSTTH 1020

OY      1245 STVATECLSKNTEENLISLKSNDLSCSNQVILAKASQEHHLSEETKCSALFSSQSELE 1304
DB      1021 STVATECLSKNTEENLISLKSNDLSCSNQVILAKASQEHHLSEETKCSALFSSQSELE 1080

OY      1305 DLTANTTODPLIGSSKQMRHOSQGVLSQDELVSDEDERGTGLEENNOEGQSDSN 1364
DB      1081 DLTANTTODPLIGSSKQMRHOSQGVLSQDELVSDEDERGTGLEENNOEGQSDSN 1140

OY      1365 L 1365
DB      1141 L 1141

RESULT 5
O46486 PRELIMINARY; PRT; 1141 AA.
ID O46486
AC O46486;
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE Breast cancer type 1 susceptibility protein (Breast and ovarian cancer
DE susceptibility protein) (Fragment).
GN BRCA1.
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98122577; PubMed=9462745;
RA Hacia J.G., Makalowski W., Edgemon K., Erdos M.R., Robbins C.M.,
RA Fodor S.P.A., Brody L.C., Collins F.S.;
RT "Evolutionary sequence comparisons using high-density oligonucleotide
RT arrays."
RL Nat. Genet. 18:155-158(1998).
CC -1- FUNCTION: NOT KNOWN, MAY REGULATE GENE EXPRESSION. MAY PLAY A ROLE
CC IN CELL PROLIFERATION AND DIFFERENTIATION (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC EMBL: AF019075; AAC39583.1; -.
DR InterPro: IPR002378; Brst_cancer1.
DR PRINTS: PR00493; BRSTCANCER1.
KM DNA-binding: Nuclear protein; Anti-oncogene.
FT NON_TER 1
FT DOMAIN 277 284 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 427 430 POLY-LYS.
FT DOMAIN 985 1006 LEUCINE-ZIPPER (POTENTIAL).
FT NON_TER 1141 1141
SQ SEQUENCE 1141 AA; 127450 MW; A4C692E5779091F1 CRC64;

Query Match      58.8%; Score 5671; DB 6; Length 1141;
Best Local Similarity 96.3%; Pred. No. 9,6e-285;
Matches 109; Conservative 11; Mismatches 31; Indels 0; Gaps 0;

OY      225 ACDEFSTDTNTEHHQPSNNDLNTTEKRAERHPEKYQSSVSNLHVEPCGNTTHASSIQ 284
DB      1 ACEFSETDVTNTEHHQPSNNDLNTTEKRAERHPEKYQSSVSNLHVEPCGNTTHASSIQ 60

```

```

QY 285 HENSLLLTDRMNVKAEFCNKSQOPGLARSQHNRMAGSKETCNDRTPTSTKKYDLNA 344
Db 61 HENSLLLTDRMNVKAEFCNKSQOPGLARSQHNRMAGSKETCNDRTPTSTKKYDLNA 120
QY 345 DPLCRKEMNOKLPCSENPRDTEVPWITLNSIQKVENMFSRSDLLGSDSDSHGSE 404
Db 121 DPLCRKEMNOKLPCSENPRDTEVPWITLNSIQKVENMFSRSDLLGSDSDSHGSE 180
QY 405 SNAKAVADVLVNEVDEYSGSSEKIDLLASDPHEALICKSEVHKSVESNIEDKIFGKT 464
Db 181 SNAKAVADVLVNEVDEYSGSSEKIDLLASDPHEALICKSEVHKSVESNIEDKIFGKT 240
QY 465 YRKRASLPNLSHVNTENLIGAFVTEPQIIOERPLTNLKRKRRTSGLHPEDFIKADLA 524
Db 241 YRKRASLPNLSHVNTENLIGAFVTEPQIIOERPLTNLKRKRRTSGLHPEDFIKADLA 300
QY 525 VQKTPEMINQSTNOTEONGOVANITNSGHNKTKDSDIONENPNPIESLEKESAFKTA 584
Db 301 VQKTPEMINQSTNOTEONGOVANITNSGHNKTKDSDIONENPNPIESLEKESAFKTA 360
QY 585 EPISSISNMELELNIHNSKAPKRNLRKRKSTRHHALELYVSRNLSPNCTELOIDSC 644
Db 361 EPISSISNMELELNIHNSKAPKRNLRKRKSTRHHALELYVSRNLSPNCTELOIDSC 420
QY 645 SSSEIKKKKKYNQMPVRHSRNLQMEGKEPATGAKSKNPNBOTSKRHSDTFPELKLTN 704
Db 421 SSSEIKKKKKYNQMPVRHSRNLQMEGKEPATGAKSKNPNBOTSKRHSDTFPELKLTN 480
QY 705 ARGSTKSNSTSELKEFNPSLPREEKEKLETYVAVSNNADPKDMLSGEVLOTERTV 764
Db 481 ARGSTKSNSTSELKEFNPSLPREEKEKLETYVAVSNNADPKDMLSGEVLOTERTV 540
QY 765 ESSSISLVPQTDYGOESISLLEVTGLKAKTEPKKCYSOCAPFNPKLJIGCSKDN 824
Db 541 ESSSISLVPQTDYGOESISLLEVTGLKAKTEPKKCYSOCAPFNPKLJIGCSKDN 600
QY 825 DTBEGFKYPLGHEVNHNSRSTSIEMESLDAOYLQNTFVYKRSQSFALFSPNMAEECAT 884
Db 601 DTBEGFKYPLGHEVNHNSRSTSIEMESLDAOYLQNTFVYKRSQSFALFSPNMAEECAT 660
QY 885 FSAHSGSLKKSOPKYTECECEKEENOGKNESIKRQVYNTAGRPVYQOKRPPDNKMC 944
Db 661 FSAHSGSLKKSOPKYTECECEKEENOGKNESIKRQVYNTAGRPVYQOKRPPDNKMC 720
QY 945 SIKGSRCLSSQFNGNETGLTPKAGLQNPYRIPLPEPKSVYTKCKNLEENP 1004
Db 721 SIKGSRCLSSQFNGNETGLTPKAGLQNPYRIPLPEPKSVYTKCKNLEENP 780
QY 1005 EHSMSPERKGNENIPSTVSTISRNINRENVFKAASSNINEVGSTNEVGSSINEIGSS 1064
Db 781 EHSMSPERKGNENIPSTVSTISRNINRENVFKAASSNINEVGSTNEVGSSINEIGSS 840
QY 1065 DENIOAELGRNGRPLNMLRLGVLOPEVYKOSLFGSNCKHPEIKQOEVEEYQVYND 1124
Db 841 DENIOAELGRNGRPLNMLRLGVLOPEVYKOSLFGSNCKHPEIKQOEVEEYQVYND 900
QY 1125 SPYLISDNLQEPOMGSHASOYCSETPDDLLDDELKEKPTSPAENDIKESAVFSKSVQK 1184
Db 901 SPYLISDNLQEPOMGSHASOYCSETPDDLLDDELKEKPTSPAENDIKESAVFSKSVQK 960
QY 1185 ELSRSPSEFTHTHLAOGYRGAKKLESSEENLSSEDEELPCFOHLLFGKVNIPQSQRH 1244
Db 961 ELSRSPSEFTHTHLAOGYRGAKKLESSEENLSSEDEELPCFOHLLFGKVNIPQSQRH 1020
QY 1245 STVATECLSKNTEENLILKNSLNDYNOVILVAKSOEHHSSEETKCSALFSSOCSELE 1304
Db 1021 STVATECLSKNTEENLILKNSLNDYNOVILVAKSOEHHSSEETKCSALFSSOCSELE 1080
QY 1305 DLTANTNTQDPEFLGSSKOMHROSEOGVGLSDKFLVSDDERGGLTENNOEBOGSDMN 1364
Db 1081 DLTANTNTQDPEFLGSSKOMHROSEOGVGLSDKFLVSDDERGGLTENNOEBOGSDMN 1140

```

```

QY 1365 L 1365
Db 1141 L 1141

RESULT 6
ID 046487 PRELIMINARY: PRT: 1140 AA.
AC 046487: Q28525;
DT 01-JUN-1998 (TRENDArel. 06, Created)
DT 01-JUN-1998 (TRENDArel. 06, Last sequence update)
DT 01-JUN-2001 (TRENDArel. 17, Last annotation update)
DE Breast cancer type 1 susceptibility protein (Breast and ovarian cancer
DE susceptibility protein) (Fragment).
GN BRCA1.
OC Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBL_Taxid=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98122577; PubMed=9462745;
RA Hacia J.G., Makalowski M., Edgemon K., Erdos M.R., Robbins C.M.,
RA Fodor S.P.A., Bredy L.C., Collins F.S.;
RT "Evolutionary sequence comparisons using high-density oligonucleotide
RT arrays."
RL Nat. Genet. 18:155-158(1998).
RN [2]
RP SEQUENCE OF 964-1028 FROM N.A.
RA Thompson M.E., Holt J.T.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: NOT KNOWN. MAY REGULATE GENE EXPRESSION. MAY PLAY A ROLE
CC IN CELL PROLIFERATION AND DIFFERENTIATION (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
DR EMBL: AF019078; AAC39586.1; -.
DR EMBL: U44730; AAB03212.1; -.
DR InterPro: IPR002378; BrstL_cancer1.
DR PRINTS: PR00493; BRSTCANCER1.
KW DNA-binding; Nuclear protein; Anti-oncogene.
FT NON_TER 1
FT DOMAIN 276 283 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 964 1005 LEUCINE-ZIPPER (POTENTIAL).
FT CONFLICT 978 978 R -> Q (in Ref. 2).
FT NON_TER 1140 1140
SO SEQUENCE 1140 AA; 126958 MW; EC18730AFA5E32E7 CRC64;

Query Match 56.4%; Score 5443.5; DB 6; Length 1140;
Best Local Similarity 92.6%; Pred. No. 5.4e-273;
Matches 1057; Conservative 35; Mismatches 47; Indels 3; Gaps 3;

```

Db 300 VOKTPEIINOQNGONQVNNITNSAHENKTKGDSIONEKNPNPTESEESAFKTKA 359  
 QY 585 EPISSISNMELNLINHSKAPKKNLRKSSRTHHALELVYSRNLSPNCELOIDSC 644  
 Db 360 EPISSISNMELNLINHSKAPKKNLRKSSRTHHALELVYSRNLSPNCELOIDSC 419  
 QY 645 SSSEELKKKKYNNOMPVRRSRNLQLMGKBPATGAKSKNPNQOTSRRHSDTFPELKLTN 704  
 Db 420 SSSEELKKKKYNNOMPVRRSRNLQLMGKBPATGAKSKNPNQOTSRRHSDTFPELKLTK 479  
 QY 705 AGCSFTKCSNTSELKFEVNPSPLPREEKEKLETVKVSNNADPKDMLSGERYLOTERSV 764  
 Db 480 VQGSFTNCNTSE-KEFVNPSLSREEKEKLETVKVSNNADPKDMLSGERYLOTERSV 538  
 QY 765 ESSSISLVPGTDYGTQOESISLLEVSTLGAKTEPNKVCQCAFAFENPKGLIHGCSNDN 824  
 Db 539 ESSSISLVPGTDYGTQOESISLLEVSTLGAKTEPNKVCQCAFAFENPKGLIHGCSNDN 598  
 QY 825 DREGFKYPLGHEVNHRSRTSIEMSEELDAQYLQNTFKVSKRSQFALFENPGNAEECAT 884  
 Db 599 DREGFKYPLGHEVNHRSRTSIEMSEELDAQYLQNTFKVSKRSQFALFENPGNAEECAT 658  
 QY 885 FSAHSGSLKKQSPKVTFECEQKEENQKNSNITKPVQVNTAGFPVYQOKDPVDNAC 944  
 Db 659 FSAHSGSLKKQSPKVTFECEQKEENQKNSNITKPVQVNTAGFPVYQOKDPVDNAC 718  
 QY 945 STKGSRFLSSOFKRNENGLITPNKHGLQNPYRPLPKSPKTKCKNLLEENR 1004  
 Db 719 STKGSRFLSSOFKRNENGLITPNKHGLQNPYRPLPKSPKTKCKNLLEENR 778  
 QY 1005 EHSMSPEREMNGEN-IPSTVSTISRNINRENVFEKASSNINEVSGSTNEVSGSINIGS 1063  
 Db 779 EHSVSPERAVGNKNIIPSTVSTISHNNIRENAFKEASSNINEVSGSTNEVSGSINIGS 838  
 QY 1064 SPENIOAELGRNRGRPLNMLRLGVLPYKOSLPESCKPEKTKQOYEBVQYVNTD 1123  
 Db 839 SPENIOAELGRNRGRPLNMLRLGVLPYKOSLPESCKPEKTKQOYEBVQYVNTD 898  
 QY 1124 FSPYILSDNLEQPMGSSHAQVCSSETPDDLDDGELKEPTSFANDIKSSAVFSKSYOK 1183  
 Db 899 FSPYILSDNLEQPMGSSHAQVCSSETPDDLDDGELKEPTSFANDIKSSAVFSKSYOK 958  
 QY 1184 GELSRSPFTHTHLAQYRGAKKLESSEENLSSEDEELPCFOHLLFGKVNNTPSQSTR 1243  
 Db 959 GELSRSPFTHTHLAQYRGAKKLESSEENLSSEDEELPCFOHLLFGKVNNTPSQSTR 1018  
 QY 1244 HSTVATECLSKNTEBULSLKNSLDCSNQVILAKASQEHHLSEETKCSGSLFSSQCSL 1303  
 Db 1019 HSTVATECLSKNTEBULSLKNSLDCSNQVILAKASQEHHLSEETKCSGSLFSSQCSL 1078  
 QY 1304 EDTANTANTQDPEFLIGSSKQMRHOSQGVGLSDKELVSDDERGTGLENNQOESQMS 1363  
 Db 1079 EDTANTANTQDPEFLIGSSKQMRHOSQGVGLSDKELVSDDERGTGLENNQOESQMS 1138  
 QY 1364 NL 1365  
 Db 1139 NL 1140

RESULT 7  
 046488 PRELIMINARY: PRT: 1146 AA.  
 AC 046488;  
 DT 01-JUN-1998 (TREMBLREL. 06, Created)  
 DT 01-JUN-1998 (TREMBLREL. 06, Last sequence update)  
 DT 01-JUN-2001 (TREMBLREL. 17, Last annotation update)  
 DE Breast cancer type 1 susceptibility protein (Breast and ovarian cancer susceptibility protein) (Fragment).  
 GN BRCAL.  
 OS Alouatta seniculus (Howler monkey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Alouattinae; Alouatta.

OX NCBI\_TaxID=9503;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-98122577; PubMed-9462745;  
 RA Hacia J.G., Makalowski W., Edgemon K., Erdos M.R., Robbins C.M., Fodor S.P.A., Brody L.C., Collins F.S.;  
 RT "Evolutionary sequence comparisons using high-density oligonucleotide arrays";  
 RL Nat. Genet. 18:155-156(1998).  
 CC -1- FUNCTION: NOT KNOWN, MAY REGULATE GENE EXPRESSION. MAY PLAY A ROLE IN CELL PROLIFERATION AND DIFFERENTIATION (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).  
 DR EMBL: AF019079; AAC39587.1; -;  
 DR InterPro: IPR002378; Brst.cancer1.  
 DR PRINTS: PR00493; BRSTCANCERT.  
 KW DNA-binding; Nuclear protein; Anti-oncogene.  
 FT NON\_TER 1  
 FT DOMAIN 278 285 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
 FT DOMAIN 200 203 POLY-SER.  
 FT DOMAIN 815 818 POLY-SER.  
 FT DOMAIN 1001 1004 POLY-GLU.  
 FT DOMAIN 996 1005 LEUCINE-ZIPPER (POTENTIAL).  
 FT NON\_TER 1146  
 SQ SEQUENCE 1146 AA; 128119 MW; BFD07EBBF851890E CRC64;

Query Match 52.0%; Score 5014.5; DB 6; Length 1146;  
 Best Local Similarity 86.2%; Pred. No. 7.9e-251;  
 Matches 991; Conservative 55; Mismatches 90; Indels 13; Gaps 7;

QY 226 CERSEEDVNTNTEHQPNSNDLNTTEKRAERHPEKYQGSVSNLHVEPCGTNTHASSLOH 285  
 Db 2 CERSEEDVNTNTEHQPNSNDLNTTEKRAERHPEKYQGSVSNLHVEPCGTNTHASSLOH 61  
 QY 286 ENSLLLTDRNVAEAEFCNKSQOPGLARSQNRRNAGSKETCNDRPTSTEEKVDLND 345  
 Db 62 ENSLLLTDRNVAEAEFCNKSQOPGLARSQNRRNAGSKETCNDRPTSTEEKVDLND 121  
 QY 346 PLCKEREMKOKLPCESENR-DTEEDVPWITLNSIOKVNMPSRSELIGSDSHDGESE 404  
 Db 122 PLHGRKEMKOKPCESENRDTEEDVPWITLNSIOKVNMPSRSELIGSDSHDGESE 181  
 QY 405 SNAKVADVLVNEVDYSGSSSEKIDLLASDPHEALICKSERVSKSVESNIEDKTFGRT 464  
 Db 182 SNAKVADVLVNEVDYSGSSSEKIDLLASDPHEALICKSERVSKSVESNIEDKTFGRT 241  
 QY 465 YKKKASLPMLSHVTEMLIGAEVTEPOIIQERPLTKLKRKRPTSGLHPDEFIKKADLA 524  
 Db 242 YKKKASLPMLSHVTEMLIGAEVTEPOIIQERPLTKLKRKRPTSGLHPDEFIKKADLA 301  
 QY 525 VOKTPEIINOQNGONQVNNITNSGHENKTKGDSIONEKNPNPTESEESAFKTKA 584  
 Db 302 VOKTPEIINOQNGONQVNNITNSGHENKTKGDSIONEKNPNPTESEESAFKTKA 360  
 QY 585 EPISSISNMELNLINHSKAPKKNLRKSSRTHHALELVYSRNLSPNCELOIDSC 644  
 Db 361 EPISSISNMELNLINHSKAPKKNLRKSSRTHHALELVYSRNLSPNCELOIDSC 420  
 QY 645 SSSEELKKKKYNNOMPVRRSRNLQLMGKBPATGAKSKNPNQOTSRRHSDTFPELKLTN 704  
 Db 421 SSSEELKKKKYNNOMPVRRSRNLQLMGKBPATGAKSKNPNQOTSRRHSDTFPELKLTN 480  
 QY 705 AGCSFTKCSNTSELKFEVNPSPLPREEKEKLETVKVSNNADPKDMLSGERYLOTERSV 764  
 Db 481 AGCSFTKCSNTSELKFEVNPSPLPREEKEKLETVKVSNNADPKDMLSGERYLOTERSV 540  
 QY 765 ESSSISLVPGTDYGTQOESISLLEVSTLGAKTEPNKVCQCAFAFENPKGLIHGCSNDN 824  
 Db 541 ESSSISLVPGTDYGTQOESISLLEVSTLGAKTEPNKVCQCAFAFENPKGLIHGCSNDN 600  
 QY 825 DREGFKYPLGHEVNHRSRTSIEMSEELDAQYLQNTFKVSKRSQFALFENPGNAEECAT 884  
 Db 601 DREGFKYPLGHEVNHRSRTSIEMSEELDAQYLQNTFKVSKRSQFALFENPGNAEECAT 660

QY	885	FSAHSGSLKAKSPVATTECECEKEENQCKNESNIKPVQTVNITAGFPVYQCKKAPVDNAC	944
Db	661	FSASCRSLKAKSPKPYTECECEKEENQCKNESNIKPVQTVNITAGFPVYQCKKAPVDNAC	720
QY	945	SINSGSRCLSSQPRGNETGLITRNKGLLQNPRIPELPPIKSFVYTKCKNMLENFE	1004
Db	721	-IEGSRCLSSQPRGNETGLITRNKGLLQNPRIPELPPIKSFVYTKCKNMLENFE	779
QY	1005	EHSMSPEREMNGEN-IPSTVSTISRNRIRENVFKEASSNI-----NEVGSSTNEVGS	1056
Db	780	EHSMSPERAMGNKNIIPSTVSTISHNN-RENAFKETSSSIYEVGSSTNEAGSSTNEVGS	838
QY	1057	SINIGSSDENIQAELGRNRPKINAMLRIGLQVPEYKQSLDSGNSCKHPEIKQETEEY	1116
Db	839	SINVGSSDENIQAELGRNRPKINAMLRIGLQVPEYKQSLDSGNSCKHPEIKQETEEY	898
QY	1117	VQVNTAFSPPLIADNLEQPMGSSHASQVCESETPDDLDDGEIKEPDSFAENDIKESSAV	1176
Db	899	VQVNTDVSILCLISYINLEQPMGSSHTSQVCESETPDDLDDGEIKEPDSFAENDIKESTY	958
QY	1177	FSKSVQKGLSRSBSPFTHTHLAGYRRGAKKLESSEENLSSEDEELPCFQHLFLGRVNN	1236
Db	959	FSKSVQKGLSRSBSPFTHTHLAGYRRGAKKLESSEENLSSEDEELPCFQHLFLGRVNN	1018
QY	1237	IPQSTSTRSYATCCLSKNTENENLSLKNLSNDCSNQVILLAKASQEHLSSEPTKCSASLP	1296
Db	1019	TP-OSTRHSSTAACTLSKNTENENLSLKSLSNDCSNQVILLAKASQEHLSSEPTKCSASLP	1077
QY	1297	SSQSESEDLTANNTQDPFLIGSSKQMRHOSQSGGLGDKELVSDDERGGLERENNO	1356
Db	1078	SSQSESEDLTANNTQDPFLIGSSKQMRHOSQSGGLGDKELVSDDERGGLERENNO	1137
QY	1357	EEQSMDSNLT 1365	
Db	1138	EEQSMDSNLT 1146	

RESULT 8

Q9ZID2 PRELIMINARY; PRT; 1812 AA.

Q9ZID2

AC 09ZID2

DT 01-MAY-1999 (TREMBLrel. 10, Created)

DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Brcal.

GN BRCAL OR BRCAL.

OS Mus musculus (Mouse).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_maxid=10090;

RN (1)

RP SEQUENCE FROM N.A.

RC STRAIN=129/SV;

RX MEDLINE=97026289; PubMed=8872468.

RA Szabo C.T., Wagner L.A., Francisco L.V., Roach J.C., Argonza R., King M.C., Ostlander E.A.; "Human, canine and murine BRCAL genes: sequence comparison among species."

RT Hum. Mol. Genet. 5:1289-1298(1996).

RT -1. SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.

CC EMBL; U68174; AAD00168.1; -

DR MGD; MGI:104537; Brcal.

DR InterPro; IPR001357; BRC1.

DR InterPro; IPR002378; Brst\_cancer1.

DR InterPro; IPR001547; GH\_5.

DR InterPro; IPR001841; Znf\_fing.

DR Pfam; PF00533; BRC1; 2.

DR Pfam; PF00097; zf-C3HC4; 1.

DR PRINTS; PRO0493; BRSTCNCERI.

DR SMART; SMO0282; BRC1; 2.

DR SMART; SMO0184; RING; 1.

DR PROSITE; PS00517; BRC1; 2.

DR PROSITE; PS00659; GLYCOSYL\_HYDROL\_F5; UNKNOWN.1.

DR	PROSITE: PS00518; ZF_RING_1; 1.
KW	Zinc-finger.
SEQUENCE	1812 AA; 198803 MW; 1A61B60230365B63 CRC64;
Query Match	50.6%; Score 4885; DB 11; Length 1812;
Best Local Similarity	56.0%; Pred. No. 7.2e-244;
Matches 1044;	Conservative 265; Mismatches 485; Indels 70; Gaps 32;
QY	1 MDLSALREVEQVNVINAMOKILECPICLLELIKEPVSTKCDHIFKCFMKLLLNQKGPQ 60
DB	1 MDLSAVQIQEYQNVNLHAMOKILECPICLLELIKEPVSTKCDHIFKCFMKLLLNQKGPQ 60
QY	61 CPLCKNDITKSLQESTRFSQLYEELKLTICAFOLDGLEFVANSNINFAKKNSEPHLKD 120
DB	61 CPLCKNEITKSLQESTRFSQLYEELRLTMAAFELDTQMLTNGEFSFKRNNSCERLNE 120
QY	121 EVSIISQMGYRRARRKLLQSEPEENSLQETSLSVQLSNIGVTRTLTKRIOPQKTSYI 180
DB	121 EASIIISQVYRRRRVRLRPOVEFGMATIKD-SLGVQLSNIGIVRSYKKNQIOTPRKKSYYI 179
QY	181 ELGSDSEEDTVNKATVYSGVDQDELQITPQGTREDSLSDSAKCAEFSFSDVTNTHNQ 240
DB	180 ELSDSSEETVTKPGDCSVRDQELLQITAPQEAQDEGKLHSADAEACEFSE-GIRNIEHHQ 238
QY	241 PSNNDLNTKEKAERHNEPKATYQGSYSNLHVEPCGNTIHASSLQHNSSLLTKDRMVE 300
DB	239 CS-DLNLPTENHATERHPEKCCOSISISNVCPECGTDAAHSLQETSLSLIEDIRMAE 297
QY	301 KAEPCNKSQOPLARSOHNRMAKSGKETCDRPTPEKKVYDLNADPLCEKREMNQILPC 360
DB	296 KAEPCNKSQOPIAVSQOSRMASKGTCDROVPSTGEKGVGNASLSDREKWTIPQSLC 357
QY	361 SENPRDTEVPVITLNNSSIQKVENFNRSDLELGSDDSHGSESNKADVDLVLNEYD 420
DB	358 PENSGATIDVPVITLNNSSVEKYNEMFNRSGEMLTGDSASARHESNAEAVALVEVNEVD 417
QY	421 EYSGSSEKIDLLADSPHEALICKSERVHKSVEYSNIEDKIFGKTYRKKASLPNLSHVTEN 480
DB	418 GGFSSSRKTLDTVPDHNHTLMCKSGADESKPVEDNISDKIFGKSYQRRKSGRPHLNHT- 476
QY	481 LIIGAFVIERPQITQERPLTNLKKRRRPTISGLHPEDFIKKALA-VQKTPREINQNTQ 539
DB	477 -IIGFETIERPQITQEOQDFTNKLKKRR--STSLQEDPFIKKASDAVQARTPDINQCTDLM 533
QY	540 EONGGVMTNNGCHENKTKGDSIQEKNKNPNPESLEKEASFKTAKEPISSTISNNELELN 599
DB	534 EPNEDAVSTTSCQENQIAGSLQKESKSHPLESLRKEPASTAGAKSISNSVSLDEVELN 593
QY	600 IHNKAPKKNRLRRKSSFRTHIALELVYSRNLSPNCTELQIDSCSSSEETIKKKRYNMP 659
DB	594 VHSKAPKKNRLRRKSSISICALPLE-PIRNSPPTCALQIDSCGSSSEETIKKNNSNO 652
QY	660 VHSNQLQMECKEPATGAKKSNKRNQESTKRKHSDTPELKLTLTNAPGSFTKCSMTSELK 719
DB	653 AGHLEPQDLETEREPAADAKK-NENENHTRKRADAPPEEKLMMKAGLLTSCSSPRKSQ 711
QY	720 EFNVSLLPREEEKLEKLEYKVSNNMEDPKDLTSGE-RVLOTREVSSESSISLVGTQY 778
DB	712 GPVNS-FQRTGTEDQLETRQMSDKAKELGDRVLGGEPSGKTTDRSEESTSVSLVDTQYD 770
QY	779 TQESTISLLEVSITLAKTPEKNVCVSOCAFENPKGLIHGCSKDNBNDETGFYKPLGHEVN 838
DB	771 TQNSVSVLDAHTVYRARTGSAQCMQFVAASENPKPLVHG-SNNAGSGTEGLKPLRLHNLN 829
QY	839 HSRETSIEMSESLDAQYLONTFKYSKQSFALNSPQNAEEECATFSNAHSGSLKKQSPK 898
DB	830 LSQE-KVMEDESELDTOYLQNTQFYSKQSFALFESKPRSPQKDC----AHSVPKSELSPK 884
QY	899 VTFEEQKEENQGNKRESNIKPYQVYNTIAGFPVYVQKQKPYMDMAKCSIKGGRFCLSSQF 958
DB	885 VYAKAKQ-ENQGGQEEFEISHQAAVAALVGPVLICQBEKLAADITCD--RSRRLLCPSSH 941
QY	959 RGNELGLITPKNHGLOMDPYRTPLEPIKTSFVTKCKRNLLEENFEHSMSPERPMGEN 1018



```

Db      942 RSENGELISATGKSGISONSFFKQSVSPIRRSIKTDNRKPLTEGRFERHTSTSEMAVGNEN 1001
QY      1019 I-PTSTSTIRNNIREVFEKESASSNINEVGSSINIGSDEDTQAELEGRNG 1077
Db      1002 ILQSTVHTVSLN-RGNACEAGS-----GSIHEVCSTGSPQGLGRNRG 1046
QY      1078 PKINAMLRVLQPEVYKOSLPKSNCKHPKIKOEVEVYQVNTFSPFLIDNLEOPM 1137
Db      1047 PKVNTVPLDLSMOPGCGQGVSPVD-KYLEIKQEGEAVC---ADPSPLFDHLEQSM 1101
QY      1138 GSSHASQVCEPDDLLDGEIKEDTSFAENDIKESAVFSKSVQKGLSRSPPTHTH 1197
Db      1102 -SGKVFQVCEPDDLLDVEIGHTSFGEGDIMERSAVNGSLRRESSRSPVTHAS 1160
QY      1198 LAGYRGAKKLESSEENLSDEDELPCHLLFGKVNINPSQSTRSHVATCLSKNTE 1257
Db      1161 KQSILHRASRKLESSESDSTEDDELPCHLL-SRISNTP-ELTRCRSAVVGIGIEKAE 1218
QY      1258 ENLLSKNSLNDSCNOVILAKASQEHHLSEETKCSASLFFSQCELEDLTANTNTODPFL 1317
Db      1219 GTQAPKMGSSSDCNNEVIMTEASQEHQFSDPRCSGRMFSQNSAAGSTANNSODSNF 1278
QY      1318 TGGSKQMRHQSESQGVGLSDKELVSDDEERGTLLENNDEQSMDSMLGRA-ASGESETP 1376
Db      1279 IPSPKQSRHQCGNEAFSLDKELISDNEEWATCLEEDNDQEE--DSIIPDEASAGYESET 1336
QY      1377 SVSEDCSGLSQSDIITTOQRQOTMNNILKLOEMLAEVLVQHSQSPNSPSTISDS 1436
Db      1337 NLSIEDC-----SOSDILITTOQRATMKYNLKLQOEMAHLEVLVQHSQSPNSPSTISD 1392
QY      1437 SALEDLRNPEQSTSEKAVLTQSKSEYPISONEGLSADFEYS-ADSSSTKKEGVER 1495
Db      1393 CALEDLPDLPEPNMNGAAILTSKINENPNVSNKASCDKFOQLHLEGPISGDESGMGR 1452
QY      1496 SPSKCPSLDDRMYMHSQSGSLQNNRNPQSOELLIKVYDVEEQLESGPHDLTETSLPR 1555
Db      1453 PSPFKSPLAGRSAGHSRHLQKRNSPQOELLQAGSE---ASSEPHNSTGQCLPR 1508
QY      1556 QDEGPRYLESGLISLSD-DPESDPEDRPRPESARNGNIPSSALKVQLKAESQSP 1614
Db      1509 RELGEGPYLGSGLISLSSRDPESESPK---EPANIGTPASTSAKIPGQVAFPSAAA 1564
QY      1615 AAHTITDLAGYNMEESEVSEKPELTASTERVNRKRMAMVSGTLPEEFMLVYFARKHH 1674
Db      1565 AGAD-----KAVGIVSKIKELTSSERADROISMVYSGTLPEEFMLVYFARKH 1617
QY      1675 TLNLITETTHVMTDAEFCERTIKYFLGIAGKRWVSYFWYQSIKERMLNHDH 1734
Db      1618 TLTDALTEETHVYIKTDAEFVCEFTLKFLGIAGKRWISYFWYVRSIGERLLVNHHE 1677
QY      1735 EVAGDVVNRNNGPKRARESDOKTFRGLETCYCPPTNMPDQLEMAVQLCGASVKE 1794
Db      1678 EVKGDVAVNRNNGPKRARESDOKTFRGLQYVCCPFTNMPKDELERMQLCGASVKE 1736
QY      1795 LSSFTLTGTHVPIVYVQPDAMTENGFAHIGQCEAPVYTRREVLDVALYQCELDY 1854
Db      1737 LPSLTHDTGAHLVYIQPSAMTSDSNCPDGLCKARLVAMVLDLSLSSYRCDLDAYL 1796
QY      1855 IPQI 1858
Db      1797 VQNI 1800

```

RESULT 9

ID 054952 PRELIMINARY; PRT; 1817 AA.

AC 054952; P97951;

DT 01-JUN-1998 (TREMBLrel. 06, Created)

DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Breast cancer type I susceptibility protein (Breast and ovarian cancer susceptibility protein).

```

GN BRCA1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RX MEDLINE=9911388; PubMed=9892727;
RA Bennett L.M., Brownlee H.A., Hagavik S., Wiseman R.W.;
RT "Sequence analysis of the rat brca1 homolog and its promoter region.";
RN Mamm. Genome 10:19-25(1999).
RL [2]
RP SEQUENCE OF 8-222 FROM N.A.
RC STRAIN=WISTAR-KYOTO; TISSUE=SPLEEN;
RX MEDLINE=96358532; PubMed=8761410;
RA Chen K.S., Shepel L.A., Haag J.D., Heil G.M., Gould M.N.;
RT "Cloning, genetic mapping and expression studies of the rat Brca1
gene.";
RL Carcinogenesis 17:1561-1566(1996).
CC -!- FUNCTION: NOT KNOWN, MAY REGULATE GENE EXPRESSION. MAY PLAY A ROLE
CC IN CELL PROLIFERATION AND DIFFERENTIATION (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC -!- SIMILARITY: CONTAINS A C3HC4-CLASS ZINC FINGER.
DR EMBL: AF036760; AAC36493.1; -.
DR EMBL: S82504; -. NOT ANNOTATED_CDS.
DR EMBL: S82502; -. NOT ANNOTATED_CDS.
DR EMBL: U060523; AAB40387.1; -.
DR EMBL: S82500; AAB37501.1; -.
DR Interpro: IPR001357; BRCT.
DR Interpro: IPR002378; Brst_cancerl.
DR Interpro: IPR001841; Znf_fing.
DR Pfam: PF00533; BRCT_2.
DR Pfam: PF00097; Zf-C3HC4; 1.
DR PRINTS: PR00493; BRSTCANCER1.
DR SMART: SM00292; BRCT; 2.
DR SMART: SM00184; RING; 1.
DR PROSITE: PS50172; BRCT; 2.
DR PROSITE: PS00518; ZF_RING_1; 1.
DR Zinc-finger: DNA-binding; Nuclear protein; Anti-oncogene.
FT ZN_FING 24 64 C3HC4-type.
FT DOMAIN 497 503 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT CONFLICT 38 38 Q -> K (IN REF. 2).
FT CONFLICT 192 192 A -> M (IN REF. 2).
SQ SEQUENCE 1817 AA; 199876 MW; COB4760F0E349A01 CMC64;

Query Match 50.5%; Score 4873.5; DB 11; Length 1817;
Best local similarity 56.2%; Pred. No. 2,9e-243;
Matches 1049; Conservative 259; Mismatches 488; Indels 69; Gaps 32;

```



OY	361	SENPDTEDEPVITLITLSSIOXKVENMFSRDELDGSDSDHGESESNKAKVADYUADLVENEV	42
Db	359	PENSATITDVPVITLITLSSIOXKVENMFSRKGEMILTSNADDRPASNABEAVALVEVNEVD	418
OY	421	EYSGSSEKIDLIASDPHEALICKSERVHSKSVESENIEDKIFCKTYRKASLPNLNHYTEN	480
Db	419	GCFFSSSKKIDLAAPPDPAVMCTSGDGEFSKPEVENIINDKIFPGTYORKSRPHLNHYTE	477
OY	481	LIIGFAVPEPOLIOBRPLNKLKRRRPRPSGHPREFIKADL-AVOKPPEMNOCNTQT	539
Db	478	-IIGFPTTEPOLIOQOPFTNKLKRRK--STCLHPEDEIKKADLTIVORISENLNOCTDQM	534
OY	540	BONGVAMVITNSGHNKTKGSDIONEKNPNPIESLEKESAFAFKTAPRISISSIMBLEJLN	599
Db	535	EPNDQAMSTISGONBRATGNDLOGRNHNPIESLRKKEPAFAPAKASINSISIDLEVEJN	594
OY	600	IHNKAPKKNRLKRRKSTRHIALELVSRNLSPPNCTELOIDSCSSSEIKKKYNOQP	659
Db	595	VHSSKAPKKNRRLRRK-STRCVLPLE-PLSRNPSPPCALOJIESCCSSSEETKKNNSNQTP	652
OY	660	VHRSNLOLMECKEAPATGAKKSNKPPEQTSKRHDSTDPPELTKLTNAPGFTKCSNISELK	719
Db	653	AGHITEPOLIETEPAAADAKK-NERPEHNRKKSASADAPFEELKMKAGLITSCSPRRKQ	711
OY	720	EFVNPDLPREKEERLETVKVSNNMADPKDLMSGB-RVLQTERSVYSSISLVPTGDXG	778
Db	712	GPVNPSS-PEBKIEIODELMQMPDNKKELDLYLAGEPSEKPEPSESTSVSLVPTQDYD	770
OY	779	TOEISLILEVSLTGAKTEPNKCVQOCQAFENPKGLIHCSCMDNRDTEGFFYPLGHEVN	838
Db	771	TONSYSLEANTVRYARTGVSQCMQFVASENPKELVHG-SNNAAGSCSCFPHPLHEJLN	829
OY	839	HSRETSIEMEESLEDAOYLQNTFKVYSKROSFALFSPNGABEECATFSAHSGLSKKOSPK	898
Db	830	HNQET-ITEMEDELDTQYLQNTFOYKSRQSPALFSLKRPQKQDCTLGARSVPSPREPK	888
OY	899	VTFPEQKNEENGKKNESNKPQYQVNIITAGFFVYQKDKPVDNAKCSIKGSRFCLSSOP	958
Db	889	VTSRGEQK-ERQGOEISEISHVOATVYVGLVPCOEKPGAVTMCA--DVSRNLCPSSHY	945
OY	959	RGNETGLITPNHGCLLOMRYRILPFLPIKFSFKTKCKMLLEENFEHMSPEREENGJN	1011
Db	946	RSCENGWLTTDXSGSIOHMFROQSVPLRSSIKTDNRKLTLEGREFPKH---ERQGNET	1000
OY	1019	-IPSTVSTISRRNNIENVFKKASSNINEVSGSSINEIGSSDENIOAELGRNG	107
Db	1003	AVOSTIHRISLNN-RGDACLEASS-----GSVLEVHSTGYNVGOGLDRNG	104
OY	1078	PKLNMLRLGYLOPEVYKOSLPGSNCKHPEIKKOEVEVQVJNDPESYLIISDNLQDM	113
Db	1048	PKVNTVSLIDSTQOPVYSKOSAPVSD-KYLEIOES-----KAVSADPSCLETSDHLEKPM	110
OY	1138	GSNSHAOVCSETPDDLDELDEKIEDTSPFANDIKESSAVFSKVSQKGLSRRSPETTH	119
Db	1102	RSDKTFQVCSETPDDLDELDEVEIQENASFOEGITEKSAIFNGSVLRESRSRSPVTHAS	116
OY	1198	LAQGRRAKKTLESSEENLSSDEDELPCFOHLLFGKVNNIIPQOSTRHSHVATBCLSKNTE	125
Db	1162	KSRSLIHRSSRKLEFSEEDSTDEDELPCFOHLL-RVSSSTP-ELTRCSSVAVQVPEREK	121
OY	1258	ENLLSLKSLNDCSNQVILAKASQHNLSSEFKCSASLPPSSOCSELEDTJATMNNODPUL	131
Db	1220	GTOAPRKSSIDBCNNNEVILGEASQYKXQSPEDAKCGSMFSQHSASALGSPANALODPPE	127
OY	1318	IGSSKQMRHQBEGVGLSDKELVSDDEBERGTGLE-NNQEBQSM--DSNLGEAASGCS	137
Db	1280	NPPSKORHQAENEAFSLDKELISDHEDMAACLEASQOEDBSIIPDS---VASGYES	133
OY	1375	ETVSSEDCGSLSSOSDILITTOORDMONHLLKLODEMALEAVLBOHQSOPNSVPSITS	143
Db	1336	FAINLSEDC-----SOSDILITTOORATIMKDLKLOODEMOLAVLBOHQSOPGSHRCPCLPA	139

QY	1435	OSASLEADJRNPOQSTSEKAVLITSOXSSEYPTISONEEGLSADRFEYS-ADSSTKNKEGV	1493
Db	1332	DPQALIEDLPDPBQNNSTGAILITSKININENPYSPQNKRACDKSOQPPDGLPSCDGKESGM	1451
QY	1494	ERSSFSKCPSLDDRMYMHMSGSLONRMYPSOEELIKVYDVEBQULEESGPHDLTETSYL	1553
Db	1452	RRPSPFKPLTSSRCSARGHSKSLONRNSTOSELLOPAXLE---KCEPHNLTRGRSCL	1507
QY	1554	PRODEGTPYLESGLSFSDDEPSPSDRPAESARVINIPTSLSALKVPLQKVAESQOS	1613
Db	1508	PRODLEGPYPBSGRLVLS---SRPDSESPKVSALVCTAPRSTALKISOGQVAGSORS	1564
QY	1614	PAAAHTTTAGYNAMEEESVSREKPEPLTASTERVNNRMSVMVSGLTPDEFMLVUYFKARKNH	1673
Db	1565	PAAG-GADT---AVVELYSKIKREPVTSPKEAREBDLSMVVSGLTPKRVMLVQKFAEKYR	1619
QY	1674	ITLNLLEETHTHVVMKTDAEFVCERTLKLYFLGILAGKMWVSYEMVYOSIKEYKMLNEND	1733
Db	1620	LALTGVITTEETHVHIKTDAAEFVCERTLKLYFLGILAGKMWISYSVWIKISIOERKILLSYHE	1679
QY	1734	FEVRBDVYVNGRHHQSPKRARESQDKIRGSLGELICGCFPTMMPDQLEBMVYOLCGASVYK	1793
Db	1660	FEVKBDVYVTSGHQSPKRSRESQK-KLEEGILQIYTCPEFTMMPKDELERRMILQLCGASVYK	1738
QY	1794	ELSSFTLGTGVHPIYVVOQPDAMTEBNGFHAIGOMCEAPVYREVVLVSVALYQOCEOLDTY	1853
Db	1739	ELPLLTTRDTGAIPIYLVQPSAMTEBNDNCPDIDGOLCKRGLVMMDWLDSISVYRCRDIDAY	1798
QY	1854	LIPQI 1858	
Db	1799	LWQNI 1803	

Query Match	Score 4386.5;	DB 6;	Length 1146;
Best Local Similarity	76.2%;	Pred. No. 2.2e-218;	
Matches 874;	Conservative 100;	Mismatches 166;	Indels 7; Gaps 6

```

Db      1 AEEFSE-DITNIEHNSCKDLNTIEKATERPEDEYSSVNLHVEPCGTHTASSIQ 59
OY      285 HENSLLTKDRMNVKAEFCNKSOGPLARSOHNWAGSKECNDRPSTEEKYDLNA 344
Db      60 HENSLLTKDRMNVKAEFCNKSOGPLARSOESWAGSKECNDRPSTEEKYDLNA 119
OY      345 DPLCEKERNKOKLPCSENPRDTEVPWTLNNSIQKVENWESRDELSDSHGSE 404
Db      120 DSGHGRKENMKRPYEPSPROTQDVPWTLNNSIQKVENWESRDELSDSHGSE 179
OY      405 SNAKVAVDLVNVEDEYSGSSEKIDLASDPHEALICKSERVHKSVESNIEDKIFGKT 464
Db      180 SNAKVALEVPNDVGYDSSEKIDLASDPHEALICKSERHSPVESNIEDKIFGKT 239
OY      465 YRRKATPLNLSHYTENLIGAFVTEPOITIOERPLTKLKRKRRTSGHLEPDIKADLA 524
Db      240 YOKKTSPLNLSHYTENLIGAFVTEPOITIOERPLTKLKRKRRTSGHLEPDIKADLA 299
OY      525 -VOKTPEMINOGTNOTQONOVANNITNSGHENKTGDSIQNEKNPPIESLEKESAFKTK 583
Db      300 VOKTPEKINOGIDQEQNDVANNITNSGHENKTGDSIQNEKNPPIESLEKESAFKTK 359
OY      584 AEPISISSIMELNLIHNSKAPKRNLRKRSSTRIHLELVYVSNLSPNCTELOIDS 643
Db      360 AEPISISSIMELNLIHNSKAPKRNLRKRSSTRIHLELVYVSNLSPNCTELOIDS 419
OY      644 CSSESEIKKKKYNOMVRSRNLQMEKREPATGAKSKSNPNQTSRDOFPPLKLT 703
Db      420 CSSESEIEKKNSSQVVRHSRKLQMEKREPATGAKSKSNPNQTSRDOFPPLKLT 479
OY      704 NAPSFTKCSNTESELEFVNPSLPREKEKELETAVVSNNAEDPKDLMSEGERVLOTES 763
Db      480 NVYVFAFNSSSKKLOEFIDPQLQREIEENLETIHVNSAKPKRLVLSGEGLQTES 539
OY      764 VESSISILVPGTDYQOESISLLEVSTICKAKTEPNKCVSQCAEFNPKGLHGSKDNK 823
Db      540 VESTISILVPGTDYQOESISLLEVSTICKAKTEPNKCVSQCAEFNPKGLHGSKDNK 599
OY      824 NDEGEFKYPLGHEVNSRFTSIEMESSELDQVLOTFVSKROSALFSPNMAEECA 883
Db      600 NDEGEFKYPLGHEVNSRFTSIEMESSELDQVLOTFVSKROSALFSPNMAEECA 659
OY      884 TFSAHSGSLKOSPKYTFEEOKEENOGKSNESNIRKQVOTVNTIAGPPVVGOK -DKPVDNA 942
Db      660 TYVAHRLRKOSPKYTFEEOKEENOGKSNESNIRKQVOTVNTIAGPPVVGOK -DKPVDNA 719
OY      943 KCSIKGSRFCSSQFRGNETGLTPNKHGLLONPYRIPPLPKSVFKCKKNLLEN 1002
Db      720 KESIKGSRFCSSQFRGNETGLTPNKHGLLONPYRIPPLPKSVFKCKKNLLEN 779
OY      1003 FFEHSHSPREKGNEN-IPSTVTSIRNNIRENVFPEASSNINEGSSITNEVGSSINEL 1061
Db      780 FFEHSHSPREKGNEN-IPSTVTSIRNNIRENVFPEASSNINEGSSITNEVGSSINEL 839
OY      1062 GSSDENIQALGRNRPKLNAMLRGLVLOPEYKOSLPSCNCKHPRIKQOEVEVQYTN 1121
Db      840 GSGENIQALGRNRPKLNAMLRGLVLOPEYKOSLPSCNCKHPRIKQOEVEVQYTN 899
OY      1122 TPFSPYLLSDNLEQPWGSHASQVCEPDDLLDGEIKEDTSPFAENDIKESSAVFSKV 1181
Db      900 TMLSTCLLSHNEQPESSHASQVCEPDDLLDGEIKEDTSPFAENDIKESSAVFSKV 959
OY      1182 OKGELSRSPPTPHHLLQGRGAKKLESSEENLSEDEELPCPHLLFGVYNNIPSS 1241
Db      960 OKGELSRSPPTPHHLLQGRGAKKLESSEENLSEDEELPCPHLLFGVYNNIPSS 1019
OY      1242 TRHSTVATECLSKNTEENLSTKNSLSDCSNOVYLLAKASQEHLSSETKCSALFSSQCS 1301
Db      1020 TRHSTVATECLSKNTEENLSTKNSLSDCSNOVYLLAKASQEHLSSETKCSALFSSQCS 1079
OY      1302 ELEDLTANTNTODPFLI--GSSKQMRHQSQGVGSLDKELVSDDERGTGLEE-NNQEE 1358

```

```

Db      1080 AVELDNTNTODPFLIMFPPSNHVRHSENOEVLSDKELVSDDERETALEBYNHEE 1139
OY      1359 OSMDSNL 1365
Db      1140 OSVDSNL 1146

RESULT 11
ID 046489 PRELIMINARY; PRT: 1141 AA.
AC 046489;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Breast cancer type 1 susceptibility protein (Breast and ovarian cancer
DE susceptibility protein) (Fragment).
GN BRCA1.
OS Galago crassicaudatus (thick-tailed galago) (Otolemur crassicaudatus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Strepsirrhini; Galagonidae; Otolemur.
OX NCBI_TaxID=9463;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98122577; PubMed=9462745;
RA Hacia J.G., Makalowski W., Edgemon K., Erdos M.R., Robbins C.M.,
RA Fodor S.P.A., Brody L.C., Collins F.S.;
RT "Evolutionary sequence comparisons using high-density oligonucleotide
RT arrays."
RL Nat. Genet. 18:155-156(1998).
CC -i- FUNCTION: NOT KNOWN, MAY REGULATE GENE EXPRESSION. MAY PLAY A ROLE
CC IN CELL PROLIFERATION AND DIFFERENTIATION (BY SIMILARITY).
CC -i- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
DR EMBL: AF019080; AAC39588.1;
DR InterPro: IPR002378; Brst_cancer1.
DR PRINTS: PR00493; BRSTCANCERT.
KW DNA-binding; Nuclear protein; Anti-oncogene.
FT DOMAIN 1
FT DOMAIN 277 284 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 199 202 POLY-SER.
FT DOMAIN 816 819 POLY-SER.
FT NON-TER 1141 1141
SQ SEQUENCE 1141 AA; 127578 MW; 2653CA894834946F CRC64;

Query Match 41.9%; Score 4040; DB 6; Length 1141;
Best Local Similarity 72.1%; Pctd. No. 1.7e-200;
Matches 831; Conservative 105; Mismatches 192; Indels 24; Gaps 15;

OY      226 CEFSEVDVNTNTEHOPSNNDLNTTERAAERHPEKYQGSVSNLHVEPCGTHTASSIQH 285
Db      2 CAYSEKDITISIEHQSNNDLNTIEKATEMHPREKYQGSVSNLHVEPCGTHTASSIQH 61
OY      286 ENSLLTKDRMNVKAEFCNKSOGPLARSOHNWAGSKECNDRPSTEEKYDLNA 345
Db      62 ESSLLTKDRMNVKAEFCNKSOGPLARSOESWAGSKECNDRPSTEEKYDLNA 121
OY      346 DPLCEKERNKOKLPCSENPRDTEVPWTLNNSIQKVENWESRDELSDSHGSE 405
Db      122 DPLCEKERNKOKLPCSENPRDTEVPWTLNNSIQKVENWESRDELSDSHGSE 181
OY      406 SNAKVAVDLVNVEDEYSGSSEKIDLASDPHEALICKSERVHKSVESNIEDKIFGKT 465
Db      182 SNAKVALEVPNDVGYDSSEKIDLASDPHEALICKSERHSPVESNIEDKIFGKT 241
OY      466 YRRKATPLNLSHYTENLIGAFVTEPOITIOERPLTKLKRKRRTSGHLEPDIKADLA 524
Db      242 YRRKATPLNLSHYTENLIGAFVTEPOITIOERPLTKLKRKRRTSGHLEPDIKADLA 301
OY      525 -VOKTPEMINOGTNOTQONOVANNITNSGHENKTGDSIQNEKNPPIESLEKESAFKTK 584
Db      302 -VOKTPEKINOGIDQEQNDVANNITNSGHENKTGDSIQNEKNPPIESLEKESAFKTK 361
OY      585 AEPISISSIMELNLIHNSKAPKRNLRKRSSTRIHLELVYVSNLSPNCTELOIDS 644

```

Db	362	EPISSTISNMKLELHNHNSKASKKRLKAKSSRIIRALVELVYNNKNSPPRHTNLQDSC	422
Oy	645	SSSEETAKKKTINQMPVHRSRNLQIMEGCEPTTGAKSKKRPMDQTSKRHSDSTPELKLITN	704
Db	422	SSSEETAKKSDQJIVHRHRKPGCLMEDEPPATGAKSKKRPMDQTSKRHSDSTPELKLITN	481
Oy	705	APGSEFTKSNSELKEFPNPSLPREKEEKA-ETPVYNNMADDEPKDLTSGERYLO--T	760
Db	482	ISSEFTKSNSELKEFPNPSLPREKEEKA-ETPVYNNMADDEPKDLTSGERYLO--T	538
Oy	761	ERSVSESSISLVPGTQDQTSISLLEVSTLGAKTEPNKCVSOGAAEFNPKGLHGCSK	820
Db	539	ERSIKTSISLVPDPDYGTQDSNLSLKVKVLRKYKTAAPNKAHSAOGTATENKEKELHGCSK	598
Oy	821	DNRNDTEFKPLGHEVHNSRETSTSEMESELDQYLOLNTFKVSRGSEFALFSPGNAAE	880
Db	599	DTGNDTEFKKPLRHEINHIOKISMEHDESLDQYLOLNTFKVSRGSEFALFSPGNAAE	655
Oy	881	ECATFSAN--SGSLKKQSPKVTFECEOEKENQGNKESNIKRPVQTVNITAGEPVVQOKDP	938
Db	656	ECATVCAQSLSASLRKGS-KVILECEQ-IENPGMKPEPIKIHQGNININTGSSVQCOKDK	713
Oy	939	VDN-AKCSIKRGSRFCSSQFRGNFTGLITPNKHGLONPRIPPLPFIKSVFTKCKKN	997
Db	714	TDDAKYSIKKASRCLSNQFRDNETESITVANKGLIOLNLYHPIPLSLRLELDTKCTN	773
Oy	998	LLEENFEHSHSPEEMGNEN-IPSTVSTISRNIRNRENVFKEASSNINEGSSSTNEVGS	1056
Db	774	LLEERFEHSHSLPEKAVGNENTVSTNTINON-RENAVKEASSSINNEVSSSTNEVGS	832
Oy	1057	SINETGSDENIQOELGNRGRPLNAMIRLGLVLOPEVYKQSLPGSNCKHPEIKQOEYEV	1116
Db	833	SVNEEGPSESEMIQOELDNRCGRPLNAVLRGLIMOPEVYKQMLPSINCEHPRIKQOEYEV	892
Oy	1117	VQVNTDPSPLISDNLEOPGSSHASOVCEPTDDLLDDEIKEDNVSFAFNDKESAV	1176
Db	893	VQVNPDPSSCLIDNLEQPRRSHASQSCSEPTDDLLVDELKENTSFANNTKESAV	952
Oy	1177	FSKSVQKGLSRSPPTHTHLLAOGYRRGAKKLESSEENLSEDEELPCFOHLLFGKYN	1236
Db	953	FSKVMRREISRSRSPPLAHILHTAOHOREVKKLESSEENMSSEDEELPCFOHLLFGKYN	1012
Oy	1237	IPSOSTRSVATGCLSKNTEENLISKNSLNDNSNOVYLAKAOENHLSFETCSAFL	1296
Db	1013	TSSOSTROSTIATECLSKTETETVLSIANTLNDNSNOVYLAKAOENHLSFETCSAFL	1072
Oy	1297	SSQGSSELDLTANTNTODPFLIGS--SKOMRHQSESOGVLSIDKELVSDDERGTGLEE	1353
Db	1073	SSQGS--EDLTANTNTODPFLMSDPLSKOMRQSENGEV-LSDNELVSNDEPREPSLEED	1129
Oy	1354	NNOFEQSDMSNL1365	
Db	1130	NHOEQSVSDSL1141	
RESULT 12			
0924EO PRELIMINARY: PRT: 1063 AA.			
AC	0924EO	01-DEC-2001 (Tremblrel. 19, Created)	
DT	01-DEC-2001	(Tremblrel. 19, Last sequence update)	
DT	01-DEC-2001	(Tremblrel. 19, Last annotation update)	
DE	BRCA1 (Fragment).		
GN	BRCA1.		
OS	Pedetes capensis (Springhaas).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Pedetidae; Pedetes.		
OX	NCBI_TaxID=10023;		
FN	SEQUENCE FROM N.A.		
EX	MEDLINE=21219190; PubMed=11319262;		
RT	Adkins R.M., Gelke B.L., Rowe D., Honeycutt R.L.;		
	"Molecular phylogeny and divergence time estimates for major rodent		

Query Match	Best Local Similarity	Score 3693;	DB 11;	Length 1063;
Matches 752;	Conservative 103;	Mismatches 203;	Indels 24;	Gaps 7;
QY	262	QGSVSNLHVEPCGCTNTHASSLQHNSSLLTLTKDRMNEKAFCKNSKOPGLARSQHRW	321	
DB	1	OGISVSNLCVEPCGCTNTHASSLLQHNSSNLTLTKDRMNEKAFCKNSKOPGLARSQHRW	60	
QY	322	AGSRTCDNRRTPTREKKVVDLADNAPLCERKENKKKLCSENPRTDEVPWITLMSIQ	381	
DB	61	AEKRTCDNRQTPSTKKKLDLADNADPYGRKKENKKKPPCSENPRTQDVSWITLMSIQ	120	
QY	382	VNEMSRSDLELGGSDSHDGESESNAKVADYDLVNEVDYSGSSEKIDLLASDPHEALI	441	
DB	121	VNEMSRSDLEMTSDDLHDGVSESNAEVGLAEVDELKHGSGSEKIDLLASDPHMLT	180	
QY	442	CKSEFVHKSVEBSNIEDKIFGKTYKKAKSLPMLSHVTENLLIGAFVTEPQIIOERPLNK	501	
DB	181	CKSEFVCAKPVESNIEDKVFQKTYRRKSLPLPLSTFTEMLLVGAATPQIOERHPPLNK	240	
QY	502	LKRKRRTPSGLHPEDFIKKADLA-V-QKTPENINQSTNTEONGGVMMNTISGHEKTKGD	560	
DB	241	LKRKRKATSPLRPEDFIKKVDLTLYQKTPERKINQSTAO MGONGVYINNINDLEMTAND	300	
QY	561	STONEKNPPIESLEKESAFKTKAEPISSSISNMELDINHNSKAPKKNRLRKSSTRI	620	
DB	301	YLKKNKNANPIESLEKESHTKAPISISISNMELDINHNSKAPKKNRLRKSSTRI	360	
QY	621	HALELYVSRNLSPMCTELOTIDSCSSSEIKKKKKKNQMPVHASNLOIMEKREPATGAK	680	
DB	361	HALELYVKNKNSPMPQETQIDSCSSSEIKKKKSDQSVKRSKRLQIMEDEHPMGAK	420	
QY	681	SNKPEQSKRHDSDTPFPELKLTLNAPGSFTKCSNLSKLEKRYVNSLPREEKEKLETYV	740	
DB	421	NNTPEKISEKHATADPEVAVKTNIPGLNNCSNKSKEEYVSPQFOKEEIEENLETQV	480	
QY	741	SNNADPMDLISGRVLOTERSVSSSSISLVPGTDYGTQESISLLEVSTLGAKEATPNK	800	
DB	481	SNSTADPNLVLASAKGLPRPERSTESTVSLVPDQDYPQOSISLLENTLTKAKATMSQ	540	
QY	801	CVSQAFAENPKGLIHGCSKDNRNDETFKXPLIGHVNHSETSIEMESELEDAQYLONT	860	
DB	541	CMQVVALAEHRKELSHSCSKDTRNTKFKDALREVNHIQETNEMESELEDAQYLONT	600	
QY	861	FKVSKROSFALFSNFGNAEECCATFSAHSGSLTKKOSPKVTEFCEQKEENQKNESNKPV	920	
DB	601	FOASKROSFALFPNSRNPEKEC---AASMSLRKSTVTLTLEYEKEENQKEESKIPKA	656	
QY	921	QTVNTTAEPVVGOKDRPVNAKKSICKGSRFLCSSQCRNGETGLTPRNKGILLONPARI	980	
DB	657	QEVNTTVAEPVVAOCQDRPCADANHSITVSRLCPQLSDLTDETELSTYNAKHGVLONPARI	716	
QY	981	PPLPEKIFSVTKCKKNLLENFEFHKSMPREEMEN-IPSTVSTISRNRIRENVFEA	1039	
DB	717	PSISPVKSFITTKKKSPSEKPEKYSASHENAMGNESITQSIYTTIQNNIRENAFREA	776	
QY	1040	SSSNINEVGSSTNEVGSSINELGSSDENIQALGNRGPKNLNMRLGLVLOPEYKOSLP	1099	
DB	777	SS-----GSSINEVGSSEMMQVGLGNRGPKNLSVYORGLHIOPEYKOSLP	822	
QY	1100	GSNCKHPEIRK-OEEVEVYVOTVNDPESYLSLDNLEQMGSHASQVSEPPDLDLGE	1158	
DB	823	VSS-KHPEMKQOENEGVQAVDADFSCPLTNSLEQPMRSGCYSQICSETPDDLDLKE	881	
QY	1159	IKETSPFENDIKSSAVFSKSVQGLSRSRSPFTTHLAQGRKAGKALLESSENLIS	1218	

Db 882 VKENTSTEGIKETSAVFSKSIQGESSSPPLTHISLAOCHORARKLESSESIS 941  
 QY 1219 EDEELPCFOHLFGKYNINIPSTOSTRSTVATECLSKNTEENLSLKNLSNDCSNQYILAK 1278  
 Db 942 EDEELPCFOHLFGKYNINIPSTOSTRSTVATECLSKNTEENLSLKNLSNDCSNQYILAK 1001  
 QY 1279 ASQEHHLSEETKCSAFLSSQCELEDTANTANTODPFLI--GSSKOMRHOSEOGVGLS 1336  
 Db 1002 TSOEPLHSEAKYSGSLFSSQCSVLEDLIPANTNTQSSFLMFGPSPKQTRHSEONEVYLS 1061  
 QY 1337 DK 1338  
 Db 1062 DK 1063

RESULT 13  
 Q8BMT3  
 ID 08BMT3 PRELIMINARY: PRT: 947 AA.  
 AC 08BMT3;  
 DT 01-MAR-2002 (TREMBlrel. 20, Created)  
 DT 01-MAR-2002 (TREMBlrel. 20, last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, last annotation update)  
 DE BRCA1 (Fragment).  
 OS Tapirus pinchaque (Mountain tapir).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Perissodactyla; Tapiridae; Tapirus.  
 OX NCBI\_TaxID=30582;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Murphy W.J., Bizik E., O'Brien S.J., Madsen O., Scally M., Donady C.,  
 RA Teeling E., Ryder O., Stanhope M.J., deJong W.W., Springer M.S.;  
 RT "Resolution of the early placental mammal radiation using Bayesian  
 RT phylogenetics.";  
 RL Science 0:0-0(2001).  
 DR EMBL; AY057830; AAL30077.1; -  
 FT NON\_TER 1  
 FT NON\_TER 947  
 FT SEQUENCE 947 AA; 104734 MW; 21D8DE08BDE0D43 CRC64;

Query Match 36.8%; Score 3557.5; DB 6; Length 947;  
 Best Local Similarity 75.1%; Pred. No. 1.1e-175;  
 Matches 713; Conservative 79; Mismatches 152; Indels 5; Gaps 5;  
 QY 274 CGTNHASSLOHENSLLTKDRMNVKAEPCNKSOGLARSHNRNAGSEKTCNDRT 333  
 Db 1 CGTNHASSLOHENSLLTKDRMNVKAEPCNKSOGLARSHNRNAGSEKTCNDRT 60  
 QY 334 PSTERKVDLADPCLERKEMNKOKLPCSENPRTDEVPWITLNSSTOKYNEVFSRSDLL 393  
 Db 61 PSEKVKLVNADPLGRKELNKQKPCSSNSRSDQDVPITLNSSTOKYNEVFSRSDLL 120  
 QY 394 GSDSDHDESESNAYADVLDVLEVDYSGSSEKIDILASPHALICKSRVHSKSYE 453  
 Db 121 SSDSDRDGPEHTEVAGAVEPNEVDGSGSEKIDILASPHALICKSERVHSKPYA 180  
 QY 454 SNIEDEKTEKTKRKAAPLNSHTVENLIGAFVPEPOLIOERPLTNKIKRRRRRTSGH 513  
 Db 181 NNEIEKITEKTKRKAAPLNSHTVENLIGAFVPEPOLIOERPLTNKIKRRRRRTSGH 240  
 QY 514 PEDFTKADLA-VQKTPKEMINGTNOTEQNGOVNMTNSGHNKTKGDSIOEKPNPPE 572  
 Db 241 PEDFTKADLA-VQKTPKEMINGTNOTEQNGOVNMTNSGHNKTKGDSIOEKPNPPE 300  
 QY 573 SLEKSAFETKAEPPISSISINMELINLHNSKAPKKNRLRRKSSRTHALELVVSRNIS 632  
 Db 301 SLEKSAFETKAEPPISSISINMELINLHNSKAPKKNRLRRKSSRTHALELVVSRNIS 360  
 QY 633 PNCGELOIDSCSSSEIKKKKKYNNOMPVHNSRLIOEKEPATGAKSKNKEQTSKH 692  
 Db 361 PNHTELQIDCCSSSEIKKKKKYNNOMPVHNSRLIOEKEPATGAKSKNKEQTSKH 420  
 QY 693 DSDTEPELKLITNAPGSEFTKCSNTSELKEFVNPSLPREEKEKLETVKYSNNAEDKDLML 752

Db 421 ASHAFPELKLITNAPGSEFTKCSNTSELKEFVNPSLPREEKEKLETVKYSNNAEDKDLML 480  
 QY 753 SGERVLQTRSVSESSISILVPGDTGTQOSISILEVSTIGRAKTEPNKCVSCAAFEENK 812  
 Db 481 SGERVLQTRSVSESSISILVPGDTGTQOSISILEVSTIGRAKTEPNKCVSCAAFEENK 540  
 QY 813 GLIHGSKDRNDTEGFKYPLGHEVNSRETSTIEMESELDOYLONTFKVSKROSFALF 872  
 Db 541 ELMHGSCKDRNDTEGFKYPLGHEVNSRETSTIEMESELDOYLONTFKVSKROSFALF 600  
 QY 873 SNPGNAEECATFSAHSGSLKQSPKVTFCQKEENQKNSNIRPVQTVNITAGFPV 932  
 Db 601 SNSGN-PEECATVCAHSAHRLRKESEPVLTLECRKEENQKESKIKHVQSVHTPAGFPV 659  
 QY 933 GQKDRPVDAKCSINGSGRFLCSSQFRGNETGLTPNKIGLLONPRIPRLPIKSFVT 992  
 Db 660 GQKDRPVDAKCSINGSGRFLCSSQFRGNETGLTPNKIGLLONPRIPRLPIKSFVT 719  
 QY 993 KCKKNLLEENFEHSMSPREKNGEN-IPSTVSTIRNNIRENVEKASSNINEVGSST 1051  
 Db 720 TCKKNVSEKLEHVSPEKRAKNGESIIOSTVSTIRNNIRENVEKASSNINEVGSST 779  
 QY 1052 NEVGSSINEGSSDENIQALGRNRPKLNMLRLGLVQPEVYKOSLPSSNCKHPEIKRO 1111  
 Db 780 NEVGSSINEGSSDENIQALGRNRPKLNMLRLGLVQPEVYKOSLPSSNCKHPEIKRO 839  
 QY 1112 -EYEVQVTVNDFSYLISDNLQDPMGSSHASQVCEFTPDLDLDEKEDTSFENDI 1170  
 Db 840 GENEQVQAVNTDFSYLISDNLQDPMGSSHASQVCEFTPDLDLDEKEDTSFENDI 898  
 QY 1171 KESSAVFSKSVQKGLSRSPFTHHLAGYRGRKTKLESSEENLSE 1219  
 Db 899 KESSAVFSKSVQKGLSRSPFTHHLAGYRGRKTKLESSEENLSE 947

RESULT 14  
 Q9BDV3  
 ID 09BDV3 PRELIMINARY: PRT: 947 AA.  
 AC 09BDV3;  
 DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, last sequence update)  
 DT 01-OCT-2001 (TREMBlrel. 18, last annotation update)  
 DE BRCA1 (Fragment).  
 OS Dicerios bicornis (Black rhinoceros).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Dicerios.  
 OX NCBI\_TaxID=9805;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=21082081; PubMed=11214318;  
 RA Madsen O., Scally M., Donady C., Kao D., DeBry R., Adkins R.,  
 RA Amrine H., Stanhope M., de Jong W., Springer M.;  
 RT "Parallel adaptive radiations in two major clades of placental  
 RT mammals.";  
 RL Nature 409:610-614(2001).  
 DR EMBL; AF284011; AAK15598.1; -  
 DR InterPro; IPR002378; Brsc\_cancerl.  
 DR PRINTS; PRO0493; BRSCCANCERL.  
 FT NON\_TER 1  
 FT NON\_TER 947  
 FT SEQUENCE 947 AA; 104753 MW; BC8158CB0F9A404B CRC64;

Query Match 36.8%; Score 3555.5; DB 6; Length 947;  
 Best Local Similarity 75.2%; Pred. No. 1.4e-175;  
 Matches 714; Conservative 86; Mismatches 143; Indels 7; Gaps 5;  
 QY 274 CGTNHASSLOHENSLLTKDRMNVKAEPCNKSOGLARSHNRNAGSEKTCNDRT 333  
 Db 1 CGTNHASSLOHENSLLTKDRMNVKAEPCNKSOGLARSHNRNAGSEKTCNDRT 60  
 QY 334 PSTERKVDLADPCLERKEMNKOKLPCSENPRTDEVPWITLNSSTOKYNEVFSRSDLL 393  
 Db 61 PSEKVKLVNADPLGRKELNKQKPCSSNSRSDQDVPITLNSSTOKYNEVFSRSDLL 120

```

Db 61 PNSEKKVLNADPLYGRRELKOKRPPCSDSPRODIDFWITRNSSIOKVNEWFSRSDTL 120
QY 394 GSDSDHDESESNKAVADVLNVEDEYSGSSEKIDLLASPHALICKSERVHSKVE 453
Db 121 TSDSDHGGPESNTEVAGAVEVNEVDEYSGSSEKIGLMAADPPALICESERVHSKVE 180
QY 454 SNIEDIKFGKTYRRKASLPNLSHTVENLIIAFTVEPOLIOERPLTNLKRKRPTSGH 513
Db 181 NNIEDKIFGKTYRRKASLPNLSHTVEDLIIIGASAIESQITQERPLTNLKRKRPTSGH 240
QY 514 PEDFIKAKDLA-VOKTPEMINOGTNOTQONQVNNITSGHENKTKGSIONEKNPPIE 572
Db 241 PEDFIKAKDLA-VOKTPEMINOGTNOTQONQVNNITSGHENKTKGSIONEKNPPIE 300
QY 573 SLEKSAFRTKAEPISSISNMNLELNHNSKAPKKNLRKRSRHHHALELVYRNLS 632
Db 301 SLEKSAFRTKAEPISSISNMNLELNHNSKAPKKNLRKRSRHHHALELVYRNLS 360
QY 633 PPNCTELOIDSCSSSEELIKKKKYNQMPVHRNRNOLMGKBPATGAKSKNPNQTSKRH 692
Db 361 PPNHTELOVDCSSSEEMKKKSPQVPRHSRKLQUTEDKBPAGAKSKNPNQJINKRL 420
QY 693 DSDFFPELKLITNAGSFTKCSNTSELKEFVNPSPPREKEBEKLETVKYSNNAEDPKDML 752
Db 421 ASDAFPELKLITNAGSFTKCSNTSELKEFVNPSPPREKEBEKLETVKYSNNAEDPKDML 480
QY 753 SGERVLOTERSVESSSISLVPGTDYGTQESISLLEVTSLGAKTEPNKCVSOCAAFENPK 812
Db 481 SGGKGLQAEERSVESGISLVPGTDYGTQESISLLEVTSLGAKTEPNKCVSOCAAFENPK 540
QY 813 GLIHGCKDNNDTDEGKTYPLGHEVNSRRETSIEMESSELDQVLONTFVYSKROSFALF 872
Db 541 ELIHDCKSDTNRNDTDEGKTYPLGHEVNSRRETSIEMESSELDQVLONTFVYSKROSFALF 600
QY 873 SNPNAEDECATFSAHSGSLKCKSPKVTFECEOKENGNKSNIKPVQVITAGFPV 932
Db 601 SNGSNEPEKCAVAHRTSLKRPBKVTLECGQKENGKESKKAHVOSVHTTVDPPV 660
QY 933 GQKD-KPVDNAKCSIKGSRFCISSQFNGNETGLITPNKHGLLONPYRIPPLEPIKSEVK 991
Db 661 CQKDKKPGDYKVCSTKESASRLQSSQFNGNETGLITPNKHGLLONPYRIPPLEPIKSEVK 720
QY 992 TKCKKNLEENFEHSHMSPEREMNGEN-IPSTVSTISNNIRRENVFEASSSNINEVSS 1050
Db 721 TICKKNVSEKLEHSHVSPERALGNKSVIOSTVSTISNNIRRENVFEASSSNINEVSS 780
QY 1051 TNEVGSSINETGSSDENIQAELGRNRPKLNAMLRLVGLQPEVYKOSLPISNCKHPEIKR 1110
Db 781 TNEVGSSINETGSSDENIQAELGRNRPKLNAMLRLVGLQPEVYKOSLPISNCKHPEIKR 840
QY 1111 Q-EYEEVAVQTVNTDFSPYLSIDNLEQPMGSSHASQVCEPDDLLDDEIKEDTSPFAEND 1169
Db 841 RGENEGVAVQTVNTDFSPYLSIDNLEQPMGSSHASQVCEPDDLLDDEIKEDTSPFAEND 900
QY 1170 IKESAVFSSVQVGEUSRSPSPFTHTHLAGQYRRGAKKLSSSENNLSSE 1219
Db 901 IKESAVFSSVQVGEUSRSPSPFTHTHLAGQYRRGAKKLSSSENNLSSE 947

```

## RESULT 15

```

Q9BDV4
ID Q9BDV4 PRELIMINARY; PRT; 942 AA.
AC Q9BDV4;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, last sequence update)
DT 01-OCT-2001 (TREMBlrel. 18, last annotation update)
DE BRCA1 (Fragment).
OS Equus caballus (Horse).
OC Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=21082081; PubMed=11214318;
RA Madgen O., Scally M., Douady C., Kao D., Debry R., Adkins R.,
RA Amrine H., Stanhope M., de Jong W., Springer M.;
RT "Parallel adaptive radiations in two major clades of placental
RT mammals";
RL Nature 409:610-614(2001).
DR EMBL; AF284010; AAK15597.1;
DR InterPro; IPR002378; Brst.CancerI.
DR PRINTS; PR00493; BRSTCANCER.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 942 AA; 104317 MW; F5C2EEELIC713C1A CRC64;

```

Query Match 36.1%; Score 3482; DB 6; Length 942;

Best Local Similarity 73.6%; Pred. No. 9e-172;

Matches 698; Conservative 90; Mismatches 151; Indels 10; Gaps 4;

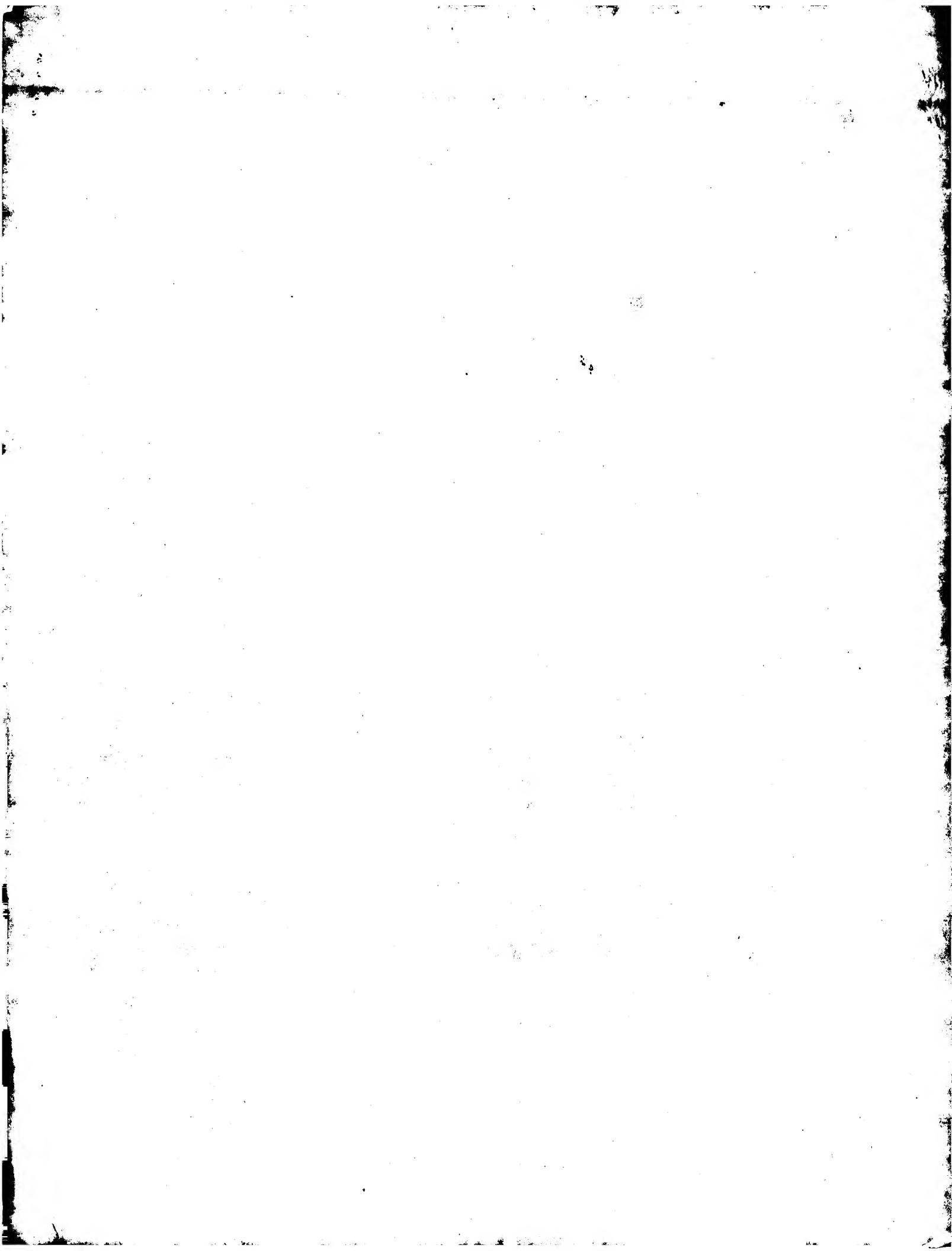
```

QY 274 CGTNTAASSLOHENSLLTYKDRANVKEAFCNKSQOPGLARSOHNRWAGSKETCNDRT 333
Db 1 CGTNTAASSLOHENSLLTYKDRANVKEAFCNKSQOPGLARSOHNRWAGSKETCNDRT 60
QY 334 PSTEKKVDLNDPLCERKEKNKOKLPCSENRDRDIEDVPWITLNSIOKVNEWFSRSDTL 393
Db 61 PNSEKKVLNADPLYGRRELKOKRPPCSDSPRODIDFWITRNSSIOKVNEWFSRSDTL 120
QY 394 GSDSDHDESESNKAVADVLNVEDEYSGSSEKIDLLASPHALICKSERVHSKVE 453
Db 121 TSDSDHGGPESNTEVAGAVEVNEVDEYSGSSEKIGLMAADPPALICESERVHSKVE 180
QY 454 SNIEDIKFGKTYRRKASLPNLSHTVENLIIAFTVEPOLIOERPLTNLKRKRPTSGH 513
Db 181 NNIEDKIFGKTYRRKASLPNLSHTVEDLIIIGASAIESQITQERPLTNLKRKRPTSGH 240
QY 514 PEDFIKAKDLA-VOKTPEMINOGTNOTQONQVNNITSGHENKTKGSIONEKNPPIE 572
Db 241 PEDFIKAKDLA-VOKTPEMINOGTNOTQONQVNNITSGHENKTKGSIONEKNPPIE 300
QY 573 SLEKSAFRTKAEPISSISNMNLELNHNSKAPKKNLRKRSRHHHALELVYRNLS 632
Db 301 SLEKSAFRTKAEPISSISNMNLELNHNSKAPKKNLRKRSRHHHALELVYRNLS 360
QY 633 PPNCTELOIDSCSSSEELIKKKKYNQMPVHRNRNOLMGKBPATGAKSKNPNQTSKRH 692
Db 361 PPNHTELOVDCSSSEEMKKKSPQVPRHSRKLQUTEDKBPAGAKSKNPNQJINKRL 420
QY 693 DSDFFPELKLITNAGSFTKCSNTSELKEFVNPSPPREKEBEKLETVKYSNNAEDPKDML 752
Db 421 ASDAFPELKLITNAGSFTKCSNTSELKEFVNPSPPREKEBEKLETVKYSNNAEDPKDML 480
QY 753 SGERVLOTERSVESSSISLVPGTDYGTQESISLLEVTSLGAKTEPNKCVSOCAAFENPK 812
Db 481 SGGKGLQAEERSVESGISLVPGTDYGTQESISLLEVTSLGAKTEPNKCVSOCAAFENPK 540
QY 813 GLIHGCKDNNDTDEGKTYPLGHEVNSRRETSIEMESSELDQVLONTFVYSKROSFALF 872
Db 541 ELIHDCKSDTNRNDTDEGKTYPLGHEVNSRRETSIEMESSELDQVLONTFVYSKROSFALF 600
QY 873 SNPNAEDECATFSAHSGSLKCKSPKVTFECEOKENGNKSNIKPVQVITAGFPV 932
Db 601 SNGSNEPEKCAVAHRTSLKRPBKVTLECGQKENGKESKKAHVOSVHTTVDPPV 660
QY 933 GQKD-KPVDNAKCSIKGSRFCISSQFNGNETGLITPNKHGLLONPYRIPPLEPIKSEVK 991
Db 661 CQKDKKPGDYKVCSTKESASRLQSSQFNGNETGLITPNKHGLLONPYRIPPLEPIKSEVK 720
QY 992 TKCKKNLEENFEHSHMSPEREMNGEN-IPSTVSTISNNIRRENVFEASSSNINEVSS 1050
Db 721 TICKKNVSEKLEHSHVSPERALGNKSVIOSTVSTISNNIRRENVFEASSSNINEVSS 780
QY 1051 TNEVGSSINETGSSDENIQAELGRNRPKLNAMLRLVGLQPEVYKOSLPISNCKHPEIKR 1110
Db 775 -NEVGSSINETGSSDENIQAELGRNRPKLNAMLRLVGLQPEVYKOSLPISNCKHPEIKR 833

```

OY 1111 QEEYVOTVNTDFSPYLTISDNLEQPMGSSHASQVCEPDDLLDDGEIKEDTSFAENDI 1170  
| : ||| | ||| | ||| | ||| | ||| | ||| : ||| :  
Db 834 QGEKEVVOAVNADFSPCLTISDNLEQPMGSSCASQVCEPDDLLNDEIKENISFAESGV 893  
OY 1171 KESSAVFSKSVOKGELSRSPFTHTHLAGYRRGAKKLESSEWLSSE 1219  
| | ||| ||| ||| : ||| | | ||| : | : ||| ||| |||  
Db 894 KERSAVFSKSVOKGELSRSPFTHTHLAGYRRGAKKLESSEWLSSE 942

Search completed: June 27, 2003, 10:26:27  
Job time : 130 secs





Page 1

PA (ONCO-) ONCORMED INC  
 22[illegible]



Db	2821	TTGAATGTGAACAAAAGGAAGAAAATCAAGGAAAGATGATGCTTAATATCAAGCCGTGTAC	288
QY	2881	AGACAGTTTAAATTCACCTGCAGGCTTTCCGTGGTTGGTCAGAAAGATTAAGCCAGTTGATA	2941
Db	2881	AGACAGTTTAAATTCACCTGCAGGCTTTCCGTGGTTGGTCAGAAAGATTAAGCCAGTTGATA	2941
QY	2941	ATGCCAAATGTAGTATCAAGAAGGCGCTAGGTTTTGTCATCATCTCAGTTCAGAGGCA	3000
Db	2941	ATGCCAAATGTAGTATCAAGAAGGCGCTAGGTTTTGTCATCATCTCAGTTCAGAGGCA	3000
QY	3001	ACGAATCTGCACCTACTTACTCCAAATTAACATGAGCTTTACAAAAACCATATTCGATAC	3066
Db	3001	ACGAATCTGCACCTACTTACTCCAAATTAACATGAGCTTTACAAAAACCATATTCGATAC	3066
QY	3061	CACCACTTTTCCATCAAGATCATTTGGTTAAACATTAATGTAAAGAAAATTCGTAGAGC	3122
Db	3061	CACCACTTTTCCATCAAGATCATTTGGTTAAACATTAATGTAAAGAAAATTCGTAGAGC	3122
QY	3121	AAACCTTTGAGGACATTCATATGTCACCTGAAGAGAAATGGGAAATGGAACATTTCCAA	3180
Db	3121	AAACCTTTGAGGACATTCATATGTCACCTGAAGAGAAATGGGAAATGGAACATTTCCAA	3180
QY	3181	GTACAGTGAACCATTTAGCCGTAATACATTAGAGAAAATGTTTTAAAGAACCGAGCT	3244
Db	3181	GTACAGTGAACCATTTAGCCGTAATACATTAGAGAAAATGTTTTAAAGAACCGAGCT	3244
QY	3241	CAAGCAATATTAAAGTAGGTTCCAGTACTAATGAAGTGGGCGCCAGTATTAAATGA	3300
Db	3241	CAAGCAATATTAAAGTAGGTTCCAGTACTAATGAAGTGGGCGCCAGTATTAAATGA	3300
QY	3301	TAGTTTCCAGTGAATAAAAACATTCACGACGACCTAGTAGAAGAACAGAGGCGCCAAATTTGA	3360
Db	3301	TAGTTTCCAGTGAATAAAAACATTCACGACGACCTAGTAGAAGAACAGAGGCGCCAAATTTGA	3360
QY	3361	ATGCTATGCTTAGATTAGGGGTTTTGGCAACCTGAGGCTCTATAACAAAGCTCTTCCTGAA	3422
Db	3361	ATGCTATGCTTAGATTAGGGGTTTTGGCAACCTGAGGCTCTATAACAAAGCTCTTCCTGAA	3422
QY	3421	GTAATTGTAAGCATTCCTGAAATTAAGAAAGCAAGATATGAACACATGCTGCAGCTGTA	3480
Db	3421	GTAATTGTAAGCATTCCTGAAATTAAGAAAGCAAGATATGAACACATGCTGCAGCTGTA	3480
QY	3481	ATACAGATTTCTCTCATATCTGATTTAGATTGATTAAGACAGCCTATGGCAACTAGTC	3540
Db	3481	ATACAGATTTCTCTCATATCTGATTTAGATTGATTAAGACAGCCTATGGCAACTAGTC	3540
QY	3541	ATGCATCTCAGGTTTGTCTCTGAGACACCTGATGACCTGTTAGATGATGGTGAATTAAGG	3600
Db	3541	ATGCATCTCAGGTTTGTCTCTGAGACACCTGATGACCTGTTAGATGATGGTGAATTAAGG	3600
QY	3601	AAGATACTAGTTTCTCTGAAAAATGCAATTAAGGAAGTCTGCTGTTTTTACCAAAACCG	3660
Db	3601	AAGATACTAGTTTCTCTGAAAAATGCAATTAAGGAAGTCTGCTGTTTTTACCAAAACCG	3660
QY	3661	TCCAGAAAGAGAGAGCTTAGACAGAGATCTTAGCCCTTACCCATATACATTTGGCTCAGG	3722
Db	3661	TCCAGAAAGAGAGAGCTTAGACAGAGATCTTAGCCCTTACCCATATACATTTGGCTCAGG	3722
QY	3721	GTTACCGAAGAGGGGCGCAAGAAATTTGAAGTCTCTGAGAGAGAACTTATCTAGTGAAGATG	3780
Db	3721	GTTACCGAAGAGGGGCGCAAGAAATTTGAAGTCTCTGAGAGAGAACTTATCTAGTGAAGATG	3780
QY	3781	AAGAGCTTCCCTGCTTCCACACACTTGTTATTTGGTAAGATTAACCATATACCTCTCAGT	3840
Db	3781	AAGAGCTTCCCTGCTTCCACACACTTGTTATTTGGTAAGATTAACCATATACCTCTCAGT	3840
QY	3841	CTACTAGGATAGACCGGTTGCTACCGAGTGTCTGTCTTAAGAACACAGAGGAGAAATTTAT	3900
Db	3841	CTACTAGGATAGACCGGTTGCTACCGAGTGTCTGTCTTAAGAACACAGAGGAGAAATTTAT	3900
QY	3901	TATCATTTGAAGAAATAGCTTAATAGTACGTCAACAGTAAATATTTGGCAAGGCACTTC	3960
Db	3901	TATCATTTGAAGAAATAGCTTAATAGTACGTCAACAGTAAATATTTGGCAAGGCACTTC	3960

QY	5041	CTTCAACAGAAAGGGTCAACAAACAAATATGCTCATGTTGTCGTGCTGACCCAGAG	5100
Db	5041	CTTCAACAGAAAGGGTCAACAAACAAATATGCTCATGTTGTCGTGCTGACCCAGAG	5100
QY	5101	AATTATGCTGCTGTACAAATTGGCAGAAAAACACACATCACTTAATCTAATTA	5160
Db	5101	AATTATGCTGCTGTACAAATTGGCAGAAAAACACACATCACTTAATCTAATTA	5160
QY	5161	CTGAAGAGACTACTCATGTTGTTATGAAAACAGATGCTGAGTTTGTGTGACGGACAC	5220
Db	5161	CTGAAGAGACTACTCATGTTGTTATGAAAACAGATGCTGAGTTTGTGTGACGGACAC	5220
QY	5221	TGAATATTTTCTAGCAATTTGGGGAGAGAAAAATGGCTAGTTAAGCATTTTGGGTGACCC	5280
Db	5221	TGAATATTTTCTAGCAATTTGGGGAGAGAAAAATGGCTAGTTAAGCATTTTGGGTGACCC	5280
QY	5281	AGTCATTATTAAGAAAGAAAAATGCTGATGAGCATGATTTTGAAGTCAGAGAGATGTGG	5340
Db	5281	AGTCATTATTAAGAAAGAAAAATGCTGATGAGCATGATTTTGAAGTCAGAGAGATGTGG	5340
QY	5341	TCAATGGAAGAAACCACCAAGGTCCTCAAGCCGAGAGAAATCCCGAGACAGAAAGATCT	5400
Db	5341	TCAATGGAAGAAACCACCAAGGTCCTCAAGCCGAGAGAAATCCCGAGACAGAAAGATCT	5400
QY	5401	TCAGGGGGCTGGAATTTGTTGCTTTGAGGCCCTTACCAACATGCCCACATCAACTGG	5460
Db	5401	TCAGGGGGCTGGAATTTGTTGCTTTGAGGCCCTTACCAACATGCCCACAGATCAACTGG	5460
QY	5461	AATGATGGTACAGCTGTGTGGTCTTGTGTGTAAGAGACTTTCATCATTCACCCCTTG	5520
Db	5461	AATGATGGTACAGCTGTGTGGTCTTGTGTGTAAGAGACTTTCATCATTCACCCCTTG	5520
QY	5521	GCACAGGTGTCACCCCAATTGTGGTTGTGACCCAGATGCTGTGACACAGACAATGGCT	5580
Db	5521	GCACAGGTGTCACCCCAATTGTGGTTGTGACCCAGATGCTGTGACACAGACAATGGCT	5580
QY	5581	TCCATGCAATTGGGAGATGTGTGAGGCACTGTGTGACCCGAGAGTGGGTGTGGACA	5640
Db	5581	TCCATGCAATTGGGAGATGTGTGAGGCACTGTGTGACCCGAGAGTGGGTGTGGACA	5640
QY	5641	GTTGTAAGCACTTACCAAGTGCAGAGAGCTGAGACCTTACCTGATACCCAGATCCCCACA	5700
Db	5641	GTTGTAAGCACTTACCAAGTGCAGAGAGCTGAGACCTTACCTGATACCCAGATCCCCACA	5700
QY	5701	GCACACTACTGA 5711	
Db	5701	GCACACTACTGA 5711	
RESULT 2			
AAC60795			
ID	AAC60795 standard; cDNA; 5711 BP.		
XX	AAC60795;		
XX	AAC60795;		
DT	07-FEB-2001 (first entry)		
XX			
DE	Human BRCA1 (om13) nucleotide sequence SEQ ID NO:5.		
XX			
KW	Human; BRCA1; chromosome 17; 17q21; breast cancer; ovarian cancer;		
KW	gene therapy; diagnosis; cytostatic; genetic susceptibility; mutation;		
XX	polymorphism; identification; ss.		
OS	Homo sapiens.		
XX			
PN	US6130322-A.		
XX			
PD	10-OCT-2000.		
XX			
PF	06-MAY-1998; 98US-0074476.		
XX			
PR	12-FEB-1996; 96US-0598591.		

PR 12-DEC-1997; 9705-0798691.

XX (GENE-) GENE LOGIC INC.

XX Zeng B, Thuber D, Olson SJ, Alvares CP, Allen ACP, Murphy PD:  
PI Critz BS;

XX WPI: 2000-646756/62.  
DR P-PSDB: AAB24219.

PT New coding sequence of the human BRCA1 gene, i.e. BRCA1 (om12), useful  
PT in gene therapy, especially for preventing or treating breast or  
PT ovarian cancer, as well as for diagnosing or monitoring breast or  
PT ovarian cancer

Example 4: Column 59-66; 56pp; English.

CC AAC60793 to AAC60795 encode the human BRCA1 (om1-3) proteins given in  
CC AAB24217 to AAB24219 respectively. BRCA1 is found on chromosome 17  
CC mapping to position 17q21. The BRCA1 (om12) coding sequence is  
CC specifically claimed in the present invention. The BRCA1 (om12) coding  
CC sequence is useful in gene therapy, especially for preventing or treating  
CC breast or ovarian cancer. It is also useful for diagnosing or monitoring  
CC breast or ovarian cancer. Furthermore, the BRCA1 (om12) coding sequence  
CC is useful for: (a) identifying individuals having BRCA1 gene mutations  
CC and having an increased genetic susceptibility to breast or ovarian  
CC cancer, or identifying a mutation that increases the genetic  
CC susceptibility to breast or ovarian cancer; (b) avoiding  
CC misinterpretation of polymorphisms found in the BRCA1 gene; (c)  
CC determining the presence of a previously unknown mutation in the BRCA1  
CC gene; (d) probing a human sample of the BRCA1 gene by allele to determine  
CC the presence of either polymorphic alleles or mutations; and (e)  
CC performing diagnosis with a reagent derived from the BRCA1 (om1) cDNA  
CC sequence. AAC60796 to AAC60861 represent PCR primers for the BRCA1 gene,  
CC which are used in an example from the present invention.

XX Sequence 5711 BP: 1956 A; 1098 C; 1274 G; 1383 T; 0 other;

Query Match 100.0%; Score 5709.4; DB 21; Length 5711;

Best Local Similarity 100.0%; Pred. No. 0;  
Matches 5710; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGCTGCTGAGACTTCTGTGACCCCGACAGGCTGTGGGTTTCTCAGATTAAGTGGCC 60  
DB 1 AGCTGCTGAGACTTCTGTGACCCCGACAGGCTGTGGGTTTCTCAGATTAAGTGGCC 60  
QY 61 CCTGCGCTCAGAGAGGCTTCCACCTGCTGGGTTAAGTTCAATGGACAGAAAGAAA 120  
DB 61 CCTGCGCTCAGAGAGGCTTCCACCTGCTGGGTTAAGTTCAATGGACAGAAAGAAA 120  
QY 121 TGAATTTATCTGCTCTTCCGCTTGAAGAGTACAAATGTCATTAATGCTATGAGAAA 180  
DB 121 TGAATTTATCTGCTCTTCCGCTTGAAGAGTACAAATGTCATTAATGCTATGAGAAA 180  
QY 181 TCTTAGAGTGTCCATCTGTGTGAGTTGATCAAGAACTGTCTCCAAAGTGTGACC 240  
DB 181 TCTTAGAGTGTCCATCTGTGTGAGTTGATCAAGAACTGTCTCCAAAGTGTGACC 240  
QY 241 ACATATTTTGCAGAAATTTTGCATGCTGCAACCTTCAACCAAGAGGCTTCCACAGT 300  
DB 241 ACATATTTTGCAGAAATTTTGCATGCTGCAACCTTCAACCAAGAGGCTTCCACAGT 300  
QY 301 GTCTTTATGTATGATGATATACCAAAAGAGCTTACAGAAAGTACGAGATTGATGTC 360  
DB 301 GTCTTTATGTATGATGATATACCAAAAGAGCTTACAGAAAGTACGAGATTGATGTC 360  
QY 361 AACTTGTGAGAGCTATGAAATCATTTGTCTTTTCACTTGACACAGAGTTTGGAGT 420  
DB 361 AACTTGTGAGAGCTATGAAATCATTTGTCTTTTCACTTGACACAGAGTTTGGAGT 420  
QY 421 ATGCAAAAGCTATTAATTTTGCAGAAAGAGAAATTAATCTCCGAGACATTAAGATG 480  
DB 421 ATGCAAAAGCTATTAATTTTGCAGAAAGAGAAATTAATCTCCGAGACATTAAGATG 480

QY 481 AAGTTTCTATCATCCAAAGTATGGCTACAGAAACCGTCCAAAGACTTCTACAGATG 540  
DB 481 AAGTTTCTATCATCCAAAGTATGGCTACAGAAACCGTCCAAAGACTTCTACAGATG 540  
QY 541 AACCAGAAATTCCTTCCCTTGCAGAGAAACCAAGTCTCAGTCCAGCTTCAACCTTGA 600  
DB 541 AACCAGAAATTCCTTCCCTTGCAGAGAAACCAAGTCTCAGTCCAGCTTCAACCTTGA 600  
QY 601 CTGTGAGAACTGTGAGGACAAAGCAGGATACCAACTTCAAAAAGAGCTGTCTACATTG 660  
DB 601 CTGTGAGAACTGTGAGGACAAAGCAGGATACCAACTTCAAAAAGAGCTGTCTACATTG 660  
QY 661 AATTGGATCTGATCTTCTGAGAGTACCGTTAATAGCACTTAATGCAATGGGGAG 720  
DB 661 AATTGGATCTGATCTTCTGAGAGTACCGTTAATAGCACTTAATGCAATGGGGAG 720  
QY 721 ATCAAGAAATTTGTAACCAATCAACCCCTCAAGAACCCAGGATGAAATCAGTTTGGATTCTG 780  
DB 721 ATCAAGAAATTTGTAACCAATCAACCCCTCAAGAACCCAGGATGAAATCAGTTTGGATTCTG 780  
QY 781 CAAAAAAGGCTGTTGTGATTTTCTGAGAGGATGTAACAATCTGAGATCATCAAC 840  
DB 781 CAAAAAAGGCTGTTGTGATTTTCTGAGAGGATGTAACAATCTGAGATCATCAAC 840  
QY 841 CCAATTAATGATTTGAACCCACTGGAAGCCGTGAGAGGATCCAGAAAGT 900  
DB 841 CCAATTAATGATTTGAACCCACTGGAAGCCGTGAGAGGATCCAGAAAGT 900  
QY 901 ATCAGAGTATGTTCTGTTTCAAACTTGCATGTGAGACCACTGTGCAAAATCTATGCA 960  
DB 901 ATCAGAGTATGTTCTGTTTCAAACTTGCATGTGAGACCACTGTGCAAAATCTATGCA 960  
QY 961 GCTCATTCACAGATGAGAACAGCACTTTATTAATGTAAGACAGAAATGATAGAA 1020  
DB 961 GCTCATTCACAGATGAGAACAGCACTTTATTAATGTAAGACAGAAATGATAGAA 1020  
QY 1021 AGGCTGAATTTCTGTAATTAAGCAACAGCCGCTTGAAGAGGACCAATCAACAT 1080  
DB 1021 AGGCTGAATTTCTGTAATTAAGCAACAGCCGCTTGAAGAGGACCAATCAACAT 1080  
QY 1081 GGGCTGGAAGTAAAGAAACATGTATGATGAGCGGACCTCCAGCAGCAAAAAAGGTAG 1140  
DB 1081 GGGCTGGAAGTAAAGAAACATGTATGATGAGCGGACCTCCAGCAGCAAAAAAGGTAG 1140  
QY 1141 ATCTGAATGCTGATCCCTGCTGTGAGAGAAAGATGGAATGAGCAAACTGCCATGCT 1200  
DB 1141 ATCTGAATGCTGATCCCTGCTGTGAGAGAAAGATGGAATGAGCAAACTGCCATGCT 1200  
QY 1201 CAGAGAAATCTAGAGATCTGAAAGATGTTCTTGAATTAACCTAATAGCAAGCATTTGCA 1260  
DB 1201 CAGAGAAATCTAGAGATCTGAAAGATGTTCTTGAATTAACCTAATAGCAAGCATTTGCA 1260  
QY 1261 AAGTTAATGAGTGTGTTCCAGAAAGTGAATGTAAGTGTGATGATGATGATGATG 1320  
DB 1261 AAGTTAATGAGTGTGTTCCAGAAAGTGAATGTAAGTGTGATGATGATGATGATG 1320  
QY 1321 GGGAGTCTGATCAAAATGCCAAGTGAATGTAATGAGTGTGATGATGATGATGATG 1380  
DB 1321 GGGAGTCTGATCAAAATGCCAAGTGAATGTAATGAGTGTGATGATGATGATGATG 1380  
QY 1381 AATATTCGTTCTTCAAGAGAAATAGACTTACAGGAGTATGATGATGATGATGATG 1440  
DB 1381 AATATTCGTTCTTCAAGAGAAATAGACTTACAGGAGTATGATGATGATGATGATG 1440  
QY 1441 TATGTAAGAGTGAAGAGTCTCACTCAATATGATGAGAGTATGATGATGATGATG 1500  
DB 1441 TATGTAAGAGTGAAGAGTCTCACTCAATATGATGAGAGTATGATGATGATGATG 1500  
QY 1501 TTGGGAAAGCTATTCGAGAGAGGCAAGCCCTCCCAACTTAAGCATGATGATGATG 1560  
DB 1501 TTGGGAAAGCTATTCGAGAGAGGCAAGCCCTCCCAACTTAAGCATGATGATGATG 1560

OY	1501	TAATATATAGGAGCATTTGTTACTAGGCCACAGATATATCAAGAGGTCCTCCACAAATA	1620
Db	1501	TAATATATAGGAGCATTTGTTACTAGGCCACAGATATATCAAGAGGTCCTCCACAAATA	1620
OY	1621	AATTAAACGCTAAAAGGAGACCTACATCAGGCTTCATCTCGAGAGATTTCACAAAGAA	1680
Db	1621	AATTAAACGCTAAAAGGAGACCTACATCAGGCTTCATCTCGAGAGATTTCACAAAGAA	1680
OY	1681	CAGATTTTGGCAGTTCAAAAAGACTCCTCGAATATGATTAATCAGGAACTAAACCAACGAGC	1740
Db	1681	CAGATTTTGGCAGTTCAAAAAGACTCCTCGAATATGATTAATCAGGAACTAAACCAACGAGC	1740
OY	1741	AGATATGCTCAAGTGTATTAATTTACTAATAGTGGTCAATGAGAGATTAACCAAAAGGTGAT	1800
Db	1741	AGATATGCTCAAGTGTATTAATTTACTAATAGTGGTCAATGAGAGATTAACCAAAAGGTGAT	1800
OY	1801	CTATTTCAGATAGAGAAAAATCCTAACCCAAATGAAATCTCGAAAAAACAATCTGGTTTCA	1860
Db	1801	CTATTTCAGATAGAGAAAAATCCTAACCCAAATGAAATCTCGAAAAAACAATCTGGTTTCA	1860
OY	1861	AAACGAAGCTGAACTATTAAGCAGCAGTATTAAGCAATATGGAACCTCGAATTAATATTC	1920
Db	1861	AAACGAAGCTGAACTATTAAGCAGCAGTATTAAGCAATATGGAACCTCGAATTAATATTC	1920
OY	1921	ACAATTTCAAAAGCACCCTTAAGAAAGATTAAGGCTGAGAGAGAAGTCTTTCACAGCATATTC	1980
Db	1921	ACAATTTCAAAAGCACCCTTAAGAAAGATTAAGGCTGAGAGAGAAGTCTTTCACAGCATATTC	1980
OY	1981	ATGCGCTTGAAGCTAGTACTAGTAGTAATCTTAAGCCCACTAATTTGCTCAATTTGCANA	2040
Db	1981	ATGCGCTTGAAGCTAGTACTAGTAGTAATCTTAAGCCCACTAATTTGCTCAATTTGCANA	2040
OY	2041	TTGATAGTTTGTCTTAGCAGTGAAGATTAAGAAAAAAAAGTACACCAATATGCCAGTCA	2100
Db	2041	TTGATAGTTTGTCTTAGCAGTGAAGATTAAGAAAAAAAAGTACACCAATATGCCAGTCA	2100
OY	2101	GGCAGCAGCAGAAACCTACAACTCATGGAAGTAAAGAACTCTGCAACTGAGCCAAAGAA	2160
Db	2101	GGCAGCAGCAGAAACCTACAACTCATGGAAGTAAAGAACTCTGCAACTGAGCCAAAGAA	2160
OY	2161	GTAACAAACCCAAATGAACAGACAAGTAAAGACATGACAGATACCTTCCACAGCGTGA	2220
Db	2161	GTAACAAACCCAAATGAACAGACAAGTAAAGACATGACAGATACCTTCCACAGCGTGA	2220
OY	2221	AGTTAAACAAATGCACTGTTCTTTTACTAAGTGTCAAAATACAGAGTAACTTTAAAGAT	2280
Db	2221	AGTTAAACAAATGCACTGTTCTTTTACTAAGTGTCAAAATACAGAGTAACTTTAAAGAT	2280
OY	2281	TTTGTCATTCCTAGCCTTCCAAAGAGAGAAAAAGAAACATGAAACAGTTAAAGTGT	2340
Db	2281	TTTGTCATTCCTAGCCTTCCAAAGAGAGAAAAAGAAACATGAAACAGTTAAAGTGT	2340
OY	2341	CTAATTAATGCTGAAGAGCCCAAGATCTCATAGTTAAGTGGAGAAAGGTTTGGAACTG	2400
Db	2341	CTAATTAATGCTGAAGAGCCCAAGATCTCATAGTTAAGTGGAGAAAGGTTTGGAACTG	2400
OY	2401	AAAGATCTGTAGAGAGTAGCAGTATTTTCATTTGTAACCTGTAAGTATATGCACTCAGG	2460
Db	2401	AAAGATCTGTAGAGAGTAGCAGTATTTTCATTTGTAACCTGTAAGTATATGCACTCAGG	2460
OY	2461	AAAGATCTGCTTACTGGAAGTATGACACTTAAGGGAAGGCAAAAAACAGAACTAATTAAT	2520
Db	2461	AAAGATCTGCTTACTGGAAGTATGACACTTAAGGGAAGGCAAAAAACAGAACTAATTAAT	2520
OY	2521	GTTGTCAGTCACTGTGCACACTTTGAAAAACCCCAAGGAGCTAATTCATGCTGTTCCAAAG	2580
Db	2521	GTTGTCAGTCACTGTGCACACTTTGAAAAACCCCAAGGAGCTAATTCATGCTGTTCCAAAG	2580
OY	2581	ATAATATAGAAATGACACAGAGAGCTTTAAGTATTCATTTGGACATGAAAGTTAACCAAGT	2640
Db	2581	ATAATATAGAAATGACACAGAGAGCTTTAAGTATTCATTTGGACATGAAAGTTAACCAAGT	2640
OY	2641	GGGAAACAAAGCATAGAAATGGAAGAAAGTGAACCTGATGCTCAGATATTTGCGAATACAT	2700

Db	2641	GGCAACAGACATACAAATGGAAGAAAGTGAACCTGATGCTCATATTTCCAGAAATCAT	2700
Oy	2701	TCAGAGTTTCAAAAGCGCCAGTCATTTGGTCTGTTTCAATATCCAGAAATGCAAGAAAGG	2760
Db	2701	TCAGAGTTTCAAAAGCGCCAGTCATTTGGTCTGTTTCAATATCCAGAAATGCAAGAAAGG	2760
Oy	2761	AATGTGCAACATCTCTGCCCCACCTCTGGGTCCTTTAAAGAAACAAAGTCCAAAAGTCATT	2820
Db	2761	AATGTGCAACATCTCTGCCCCACCTCTGGGTCCTTTAAAGAAACAAAGTCCAAAAGTCATT	2820
Oy	2821	TTGAATGTGACAAAGGAAGAAATATCAAGGAAGAAATGAGTCTAATATTCAGCTGTATC	2880
Db	2821	TTGAATGTGACAAAGGAAGAAATATCAAGGAAGAAATGAGTCTAATATTCAGCTGTATC	2880
Oy	2881	AGACAGTTAATATCATCTGACGCTCTTCCTGCTGGTGTGTCAGAAAGATTAAGCCAGTTGATA	2940
Db	2881	AGACAGTTAATATCATCTGACGCTCTTCCTGCTGGTGTGTCAGAAAGATTAAGCCAGTTGATA	2940
Oy	2941	ATGCCAAATGTATATCAAGAGGCTCTAGGTTTGTCTATCATCTTCACGTTCCAGAGCA	3000
Db	2941	ATGCCAAATGTATATCAAGAGGCTCTAGGTTTGTCTATCATCTTCACGTTCCAGAGCA	3000
Oy	3001	ACGAACCTGAGCTCATTACCTCAATATCAAGAGGCTCTAGGTTTGTCTATCATCTTCAGAGCA	3060
Db	3001	ACGAACCTGAGCTCATTACCTCAATATCAAGAGGCTCTAGGTTTGTCTATCATCTTCAGAGCA	3060
Oy	3061	CACCACTTTTCCATCAAGTCAATTTGTTAAACATAATGTAAAGAAATCTGCTAGAGG	3120
Db	3061	CACCACTTTTCCATCAAGTCAATTTGTTAAACATAATGTAAAGAAATCTGCTAGAGG	3120
Oy	3121	AAAACCTTTGAGAACATCTCAATGTACCTGAAAGAGAAATGGAACATCTTCACAA	3180
Db	3121	AAAACCTTTGAGAACATCTCAATGTACCTGAAAGAGAAATGGAACATCTTCACAA	3180
Oy	3181	GTACAGTATGACACATTTAGCCGTATATACATTAAGAGAAATGTTTAAAGAAAGCCAGCT	3240
Db	3181	GTACAGTATGACACATTTAGCCGTATATACATTAAGAGAAATGTTTAAAGAAAGCCAGCT	3240
Oy	3241	CAGCAATATTTAATGAAGTAGTCTGCTCAGTACTAATGAAGAGGCTCCAGTATTAATGA	3300
Db	3241	CAGCAATATTTAATGAAGTAGTCTGCTCAGTACTAATGAAGAGGCTCCAGTATTAATGA	3300
Oy	3301	TAGGTTCCAGTATGAAAAACATTCAGCAGAACTAGGTAGAAACAGAGGCGCCAAATATGA	3360
Db	3301	TAGGTTCCAGTATGAAAAACATTCAGCAGAACTAGGTAGAAACAGAGGCGCCAAATATGA	3360
Oy	3361	ATGCTATCTTAGATTTAGGGGTTTTGCAACCTGAGGCTATTAACAAAGTCTTCTGGAA	3420
Db	3361	ATGCTATCTTAGATTTAGGGGTTTTGCAACCTGAGGCTATTAACAAAGTCTTCTGGAA	3420
Oy	3421	GTAATGTAAACATCTCGAAATTAAGAAAGCAAGAAATGTGAAGAAATGTGTTACACAGTTTA	3480
Db	3421	GTAATGTAAACATCTCGAAATTAAGAAAGCAAGAAATGTGAAGAAATGTGTTACACAGTTTA	3480
Oy	3481	ATACAGATTTCTTCACATATCTGATTTACAGTAACCTTAGAACAAGCCTATGGAAGTATGTC	3540
Db	3481	ATACAGATTTCTTCACATATCTGATTTACAGTAACCTTAGAACAAGCCTATGGAAGTATGTC	3540
Oy	3541	ATGCATCTCAGGTTTGTCTGAGACACCTGATGACTGTTAGATGATGTTGAATTAAGG	3600
Db	3541	ATGCATCTCAGGTTTGTCTGAGACACCTGATGACTGTTAGATGATGTTGAATTAAGG	3600
Oy	3601	AAGATACTAGTTTGTGGAAGAAATGACATTAAGAGAAATCTGCGTGTTTTACGAAAGGG	3660
Db	3601	AAGATACTAGTTTGTGGAAGAAATGACATTAAGAGAAATCTGCGTGTTTTACGAAAGGG	3660
Oy	3661	TTCCAGAAAGAGACCTTAGCAGAGAGTCTTAGCCCTTACACCATACACATTTGGCTCAGG	3720
Db	3661	TTCCAGAAAGAGACCTTAGCAGAGAGTCTTAGCCCTTACACCATACACATTTGGCTCAGG	3720
Oy	3721	GTTACCGAAGAGGGGCCAAGAAATTTAGAGTCTTCAGAAAGAACTTATCTAGTGAGATG	3780
Db	3721	GTTACCGAAGAGGGGCCAAGAAATTTAGAGTCTTCAGAAAGAACTTATCTAGTGAGATG	3780



Db	3721	GTATCCGAAGAGGGGGCCAAAGAAATTAGAGTCTTAGAGAGAACTTATCTAGTAGAGATG	3780
Oy	3781	AAGAGCTTCCCTGCTCCACACACTTGTTATTATTGGTAAAGTAACAAATATACCTTCTCAGT	3840
Db	3781	AAGAGCTTCCCTGCTCCACACACTTGTTATTATTGGTAAAGTAACAAATATACCTTCTCAGT	3840
Oy	3841	CTACTAGGCATAGCACCGTGTGCTACCGAGTGTCTGTCTAAGACACAGAGAGAAATTTAT	3900
Db	3841	CTACTAGGCATAGCACCGTGTGCTACCGAGTGTCTGTCTAAGACACAGAGAGAAATTTAT	3900
Oy	3901	TATCATTTGAACAAATAGCTTAAATATACGTACAGTAACTCCGGTAATTTTGGCAAAAGCATCTC	3960
Db	3901	TATCATTTGAACAAATAGCTTAAATATACGTACAGTAACTCCGGTAATTTTGGCAAAAGCATCTC	3960
Oy	3961	AGGAACATACACCTTAGTGGAGGAAACAAATTTCTGCTAGCTGTATTCTTCAACAGTCA	4020
Db	3961	AGGAACATACACCTTAGTGGAGGAAACAAATTTCTGCTAGCTGTATTCTTCAACAGTCA	4020
Oy	4021	GTGAATTGGAAAGACTTGACTGCCAATATCAAAACACCCAGAGATCCTTCTTGATTGGTTCTT	4080
Db	4021	GTGAATTGGAAAGACTTGACTGCCAATATCAAAACACCCAGAGATCCTTCTTGATTGGTTCTT	4080
Oy	4081	CCAAACAATAGAGCATAGTCTGAANCCAGGAGTGTGCTGTAGAGTACAGAGAAATTTGG	4140
Db	4081	CCAAACAATAGAGCATAGTCTGAANCCAGGAGTGTGCTGTAGAGTACAGAGAAATTTGG	4140
Oy	4141	TTTCAGATGATGAAGAAAGAGAGAAACGGGCTTGGAAAGAAATTAATCAAGAAAGCAAAACA	4200
Db	4141	TTTCAGATGATGAAGAAAGAGAGAAACGGGCTTGGAAAGAAATTAATCAAGAAAGCAAAACA	4200
Oy	4201	TGATATCAACTTAAGTGAAGACAGCATCTGGGTGTGAAGTGAACAAGCGCTCTGTAAG	4260
Db	4201	TGATATCAACTTAAGTGAAGACAGCATCTGGGTGTGAAGTGAACAAGCGCTCTGTAAG	4260
Oy	4261	ACTGTCGGGCTATCCTCTCAGAGTGAATTTTAAACCATCTCAGACAGAGGATACCATGC	4320
Db	4261	ACTGTCGGGCTATCCTCTCAGAGTGAATTTTAAACCATCTCAGACAGAGGATACCATGC	4320
Oy	4321	AACATTAACCTATTAAGAGTCCAGAGAGAAATGGCTGAACCTAGAGCTGTGTAGAACAGC	4380
Db	4321	AACATTAACCTATTAAGAGTCCAGAGAGAAATGGCTGAACCTAGAGCTGTGTAGAACAGC	4380
Oy	4381	ATGGAGCCAGCCCTTCTAACAGCTAACCCCTTTCATTAAGTGAAGTCTTCTGCCCCGTGAGG	4440
Db	4381	ATGGAGCCAGCCCTTCTAACAGCTAACCCCTTTCATTAAGTGAAGTCTTCTGCCCCGTGAGG	4440
Oy	4441	ACCTGCGCAATCCAGAACACAAAGACATTCAGAAAAAGAGATTTAACTTCCACAGAAAAATA	4500
Db	4441	ACCTGCGCAATCCAGAACACAAAGACATTCAGAAAAAGAGATTTAACTTCCACAGAAAAATA	4500
Oy	4501	GTGAATACCTATTAAGCCAGAAATCCAGAAAGCCCTTTCGCTGACAAGTTTGGAGTGTCTG	4560
Db	4501	GTGAATACCTATTAAGCCAGAAATCCAGAAAGCCCTTTCGCTGACAAGTTTGGAGTGTCTG	4560
Oy	4561	CAGATAGTTCTACAGTAAAAATTAAGAACCCAGAGTGGAAAGGTCAATCCCTTCTTAAT	4620
Db	4561	CAGATAGTTCTACAGTAAAAATTAAGAACCCAGAGTGGAAAGGTCAATCCCTTCTTAAT	4620
Oy	4621	GCCCATCATTTAGATGATAGGTGGTACATGCACAGTGTCTGCGAGATCTTCAGATAGAA	4680
Db	4621	GCCCATCATTTAGATGATAGGTGGTACATGCACAGTGTCTGCGAGATCTTCAGATAGAA	4680
Oy	4681	ACTACCCATCCAGAGAGAGCTCAATTAAAGTGTGTGAATGTGAGAGACCAACAGCTGGAG	4740
Db	4681	ACTACCCATCCAGAGAGAGCTCAATTAAAGTGTGTGAATGTGAGAGACCAACAGCTGGAG	4740
Oy	4741	AGTCTGGGCCACAGATTTGAACGAAACATTTACTTGTGCCAAGCAGATCTTAAGAGGAA	4800
Db	4741	AGTCTGGGCCACAGATTTGAACGAAACATTTACTTGTGCCAAGCAGATCTTAAGAGGAA	4800
Oy	4801	CCCCCTACTGTGAATCTGGAATCAAGCCTCTTCTGTGATAGACCTGGAATCTGATCTTCTG	4860
Db	4801	CCCCCTACTGTGAATCTGGAATCAAGCCTCTTCTGTGATAGACCTGGAATCTGATCTTCTG	4860

QY	4861	AAGACAGAGCCCCAGAGTCAGCTGCTGTTGGCAGACATACCATCTTCAACCTCTGCATTTGA	49220
Db	4861	AAGACAGAGCCCCAGAGTCAGCTGCTGTTGGCAGACATACCATCTTCAACCTCTGCATTTGA	49220
QY	4921	AAGTTCGCCCAATTGGAAGTGCAGAAATCTGGCCAGATCCAGCTGCTGCTATACTACTG	49808
Db	4921	AAGTTCGCCCAATTGGAAGTGCAGAAATCTGGCCAGATCCAGCTGCTGCTATACTACTG	49808
QY	4981	ATACTGCTGGGTATTAATGCAATGGAAGAAGTGTGAGCAGGAGAGAAGCCAGAAATTGCAG	50400
Db	4981	ATACTGCTGGGTATTAATGCAATGGAAGAAGTGTGAGCAGGAGAGAAGCCAGAAATTGCAG	50400
QY	5041	CTTCAACAGAAAGGCTCAACAAGAAATGTCCATGSGTGTCTTGCCCTGACCCAGAG	51000
Db	5041	CTTCAACAGAAAGGCTCAACAAGAAATGTCCATGSGTGTCTTGCCCTGACCCAGAG	51000
QY	5101	AATTATCTGCTGTACGAAGTTTGGCAGAAAACACACATACCTTAACTAATTAATTA	51600
Db	5101	AATTATCTGCTGTGTACGAAGTTTGGCAGAAAACACACATACCTTAACTAATTAATTA	51600
QY	5161	CTGAAGAGACTACTCATCTGTTATTGAAGAACAGATGCTGAGTTGTGTGTAAGGACAC	52200
Db	5161	CTGAAGAGACTACTCATCTGTTATTGAAGAACAGATGCTGAGTTGTGTGTAAGGACAC	52200
QY	5221	TGAATATTTTCTTAGAGATTCGGGAGAGAAAATGGGTAGTTAGCTATTTCTGGGTGACC	52800
Db	5221	TGAATATTTTCTTAGAGATTCGGGAGAGAAAATGGGTAGTTAGCTATTTCTGGGTGACC	52800
QY	5281	AGTCATTTAAAGAAAGAAATGCTGATGAGCATGATTTTGAAGTCAGAGAGATGAG	53400
Db	5281	AGTCATTTAAAGAAAGAAATGCTGATGAGCATGATTTTGAAGTCAGAGAGATGAG	53400
QY	5341	TCAATGAGAGAAACCCAGAGTCCAAAGCGAGCAAGAGATCCAGAGCAAGAAAGATCT	54000
Db	5341	TCAATGAGAGAAACCCAGAGTCCAAAGCGAGCAAGAGATCCAGAGCAAGAAAGATCT	54000
QY	5401	TCAGGGGCTTGAATCTGTTGCTATGAGGCGCCTTACCACATGCGCCACAGATCAACTGG	54600
Db	5401	TCAGGGGCTTGAATCTGTTGCTATGAGGCGCCTTACCACATGCGCCACAGATCAACTGG	54600
QY	5461	AATGATGATGATGAGCTGTGTGCTTCTGTGTTGAAGAGACTTTTCATCTTACCCCTTG	55200
Db	5461	AATGATGATGATGAGCTGTGTGCTTCTGTGTTGAAGAGACTTTTCATCTTACCCCTTG	55200
QY	5521	GCACAGGTGTCACCCAAATTGTTGTTGTGACCCAGATGCTTGAACAGAGCAATGGCT	55800
Db	5521	GCACAGGTGTCACCCAAATTGTTGTTGTGACCCAGATGCTTGAACAGAGCAATGGCT	55800
QY	5581	TCCATGCAATTTGGGAGATGTGTGAGGAGCCCTGTGTGATACCCGAGAGGGGTGTTGACA	56400
Db	5581	TCCATGCAATTTGGGAGATGTGTGAGGAGCCCTGTGTGATACCCGAGAGGGGTGTTGACA	56400
QY	5641	GTGTAGCACTGTACAGATGCCAGAGAGCTGACACACTACCTGATACCCAGATGCCCCACA	57000
Db	5641	GTGTAGCACTGTACAGATGCCAGAGAGCTGACACACTACCTGATACCCAGATGCCCCACA	57000
QY	5701	GCCACTACTGA 5711	
Db	5701	GCCACTACTGA 5711	
RESULT 3			
AA17493			
AA17493 standard; cDNA: 5914 BP.			
AA17493;			
02-OCT-1996 (first entry)			
Mutated BRCA1 coding sequence from PM06.			
Cancer therapy: breast and ovarian cancer predisposing gene; Immunogen:			



KW antibody production; germline alteration; probe; lesion neoplasia; human;  
 KW gene therapy; protein replacement therapy; protein mimetic; BRCA1; ds.  
 XX Homo sapiens.  
 FH Key Location/Qualifiers  
 FT CDS 120..5711  
 FT /\*tag= a  
 FT /product= BRCA1 mutant  
 FT 2731  
 FT mutation /note= b  
 FT /\*tag= b  
 FT /note= "C to T mutation"  
 XX  
 XX PN WO9605306-A2.  
 XX PD 22-FEB-1996.  
 XX XX 11-AUG-1995: 95MO-US10202.  
 XX PR 07-JUN-1995: 95US-0483553.  
 PR 12-AUG-1994: 94US-0289221.  
 PR 02-SEP-1994: 94US-0300266.  
 PR 16-SEP-1994: 94US-0308104.  
 PR 29-NOV-1994: 94US-0348824.  
 PR 24-MAR-1995: 95US-0409305.  
 PR 07-JUN-1995: 95US-0480784.  
 XX  
 PA (MYRI-) MYRIAD GENETICS INC.  
 PA (CANC-) CANCER INST.  
 PA (RECH-) CENT RECH DU CHUL.  
 XX  
 PI Durocher F, Emi M, Nakamura Y, Simard J, Shattuck-Eidens DM;  
 DR WPI: 1996-139702/14.  
 DR P-PSDB: AAR81535.  
 XX  
 PT New nucleic acid and polypeptide for mutant or polymorphic BRCA1  
 PT gene - for diagnosis and therapy of human breast and ovarian cancer  
 PT and for diagnosing pre-disposition to these cancers  
 PS  
 PS Claim 1: ; 218pp; English.

AA17439-T17453 and AA17455-T17529 represent mutations of the human  
 breast and ovarian cancer predisposing gene (BRCA1) (for wild type see  
 CC AA17448). Proteins encoded by these mutations (see AAR81483-R81497 and  
 CC AAR81499-R81546) can be used as immunogens for antibody production.  
 CC These mutant BRCA1 genes have at least 1 mutation or polymorphism in  
 CC comparison to the wild type sequence. By detecting a germline alteration  
 CC in the wild type BRCA1 gene, a predisposition for breast and ovarian  
 CC cancer can be diagnosed. In one method, BRCA1 mRNA isolated from a tissue  
 CC sample from a subject has a probe, corresponding to a fragment of the  
 CC wild type sequence (or an allele-specific probe for one of these  
 CC mutations), added to it. The conditions allow for hybridisation of the  
 CC probe to the mRNA, and any hybridisation which occurs is detected.  
 CC Alternatively the BRCA1 gene in the tissue sample is isolated, and a  
 CC shift in electrophoretic mobility of single stranded DNA from the sample  
 CC on a non-denaturing polyacrylamide gel indicates a mutation. These  
 CC methods of detection can also diagnose a lesion neoplasia associated with  
 CC the BRCA1 locus. The methods may be used in gene therapy, protein  
 CC replacement therapy and protein mimetics, and may be used to screen for  
 CC drugs in cancer therapy.

Sequence 5914 BP; 2006 A; 1155 C; 1316 G; 1437 T; 0 other;

Query Match 100.0%; Score 5709.4; DB 17; Length 5914;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 5710; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGCTGCTGAGACTTCTGAGACCCGACACAGGCTGTGGGTTTCTCAGATACTGGCC 60.  
 DB 1 AGCTGCTGAGACTTCTGAGACCCGACACAGGCTGTGGGTTTCTCAGATACTGGCC 60  
 QY 61 CCGCGCTCAGAGAGCCCTTCAACCTCTGCTGTGGTAAAGTTCAATTGGAAGAAAGAAA 120

DB 61 CCGCGCTCAGAGAGCCCTTCAACCTCTGCTGTGGTAAAGTTCAATTGGAAGAAAGAAA 120  
 QY 121 TGGATTTATCTGCTTCGGGTGAGACGTAACAAATGCTATATGATATGCAAGAAA 180  
 DB 121 TGGATTTATCTGCTTCGGGTGAGACGTAACAAATGCTATATGATATGCAAGAAA 180  
 QY 181 TCTTAGAGTGTCCATCTGTGTGAGTGTATGATGAGACCTGTCTCCAAAGTGTGACC 240  
 DB 181 TCTTAGAGTGTCCATCTGTGTGAGTGTATGATGAGACCTGTCTCCAAAGTGTGACC 240  
 QY 241 ACATATTTTGGCAATTTTGCATGCTGAACCTCTCAACGAGAAAGGGCCTTCACAGT 300  
 DB 241 ACATATTTTGGCAATTTTGCATGCTGAACCTCTCAACGAGAAAGGGCCTTCACAGT 300  
 QY 301 GTCTTTATGTAAGATGATATACCAAGAGAGCCTCAAGAAAGTACGAGATTTAGTC 360  
 DB 301 GTCTTTATGTAAGATGATATACCAAGAGAGCCTCAAGAAAGTACGAGATTTAGTC 360  
 QY 361 AACTGTGTAAGAGCTATTTGAAATATTTGCTTTTCAGCTTGACACAGGTTTGAGT 420  
 DB 361 AACTGTGTAAGAGCTATTTGAAATATTTGCTTTTCAGCTTGACACAGGTTTGAGT 420  
 QY 421 ATGCAAAACGCTATATTTTGCAGAAAAGGAAATTAACCTCTCTGACATCTTAAAGATG 480  
 DB 421 ATGCAAAACGCTATATTTTGCAGAAAAGGAAATTAACCTCTCTGACATCTTAAAGATG 480  
 QY 481 AAGTTCTATCATCCAAATATGGGCTACAGAAACCGGCTCAAAAGACTCTTACAGATG 540  
 DB 481 AAGTTCTATCATCCAAATATGGGCTACAGAAACCGGCTCAAAAGACTCTTACAGATG 540  
 QY 541 AACCAGAAATCCCTTCTGAGAAACCCAGTCTAGTCTCACTCTAATCTTGAAGT 600  
 DB 541 AACCAGAAATCCCTTCTGAGAAACCCAGTCTAGTCTCACTCTAATCTTGAAGT 600  
 QY 601 CTGTGGAACCTGAGAGCAAAAGCAGGATATCAACTCAAAAGCGTCTCTACATG 660  
 DB 601 CTGTGGAACCTGAGAGCAAAAGCAGGATATCAACTCAAAAGCGTCTCTACATG 660  
 QY 661 AATGGGATCTGATTTCTGTAAGATACCGTAAATAGGCAACTATTTGGCGGAG 720  
 DB 661 AATGGGATCTGATTTCTGTAAGATACCGTAAATAGGCAACTATTTGGCGGAG 720  
 QY 721 ATCAAGAAATTTTACAAATCACCCCTCAAGAACCGAGGATCAATCAATGTTGGATCTG 780  
 DB 721 ATCAAGAAATTTTACAAATCACCCCTCAAGAACCGAGGATCAATCAATGTTGGATCTG 780  
 QY 781 CAAAAAGGCTCTGTGATTTTCTGAGAGGATGTACAAATTAATCAATCATCAAC 840  
 DB 781 CAAAAAGGCTCTGTGATTTTCTGAGAGGATGTACAAATTAATCAATCATCAAC 840  
 QY 841 CCAGTAATATGATTTTGAACACCACTGAGAAAGCGTGAAGGATCCACAAAAGT 900  
 DB 841 CCAGTAATATGATTTTGAACACCACTGAGAAAGCGTGAAGGATCCACAAAAGT 900  
 QY 901 ATCAGAGTAGTCTCTTTCAAACTTGATGTGAGACCATGTGGCAAAATTAATCAATG 960  
 DB 901 ATCAGAGTAGTCTCTTTCAAACTTGATGTGAGACCATGTGGCAAAATTAATCAATG 960  
 QY 961 GCTCATTACAGATGAGACAGCACTTTATTAATCAATTAAGACAGATGATAGAAA 1020  
 DB 961 GCTCATTACAGATGAGACAGCACTTTATTAATCAATTAAGACAGATGATAGAAA 1020  
 QY 1021 AGGCTGAATTCGTATATTAAGCAACAGCCTGCTTAGCAAGAGGCAACATPAACGAT 1080  
 DB 1021 AGGCTGAATTCGTATATTAAGCAACAGCCTGCTTAGCAAGAGGCAACATPAACGAT 1080  
 QY 1081 GGGCTGGAAGTAAGCAACATGATATGATAGCGGAGTCCAGAGCAGAAAGAGTAG 1140  
 DB 1081 GGGCTGGAAGTAAGCAACATGATATGATAGCGGAGTCCAGAGCAGAAAGAGTAG 1140  
 QY 1141 ATCTGAATGCTGATCCCTGTGTGAGAGAAAGATGAATTAACAGAAACTGCGATGCT 1200  
 DB 1141 ATCTGAATGCTGATCCCTGTGTGAGAGAAAGATGAATTAACAGAAACTGCGATGCT 1200

```
Db 1141 ATCTGAATGCTGATCCCTGTGTGAGAGAAAAAGTAATGAAGCAAAACTGCCATGCT 1200
OY 1201 CAGAGAACTCTAGATACCTGAAGATGTTCTTGATTAACACTAAATAGCAGCATTCAGA 1260
Db 1201 CAGAGAACTCTAGATACCTGAAGATGTTCTTGATTAACACTAAATAGCAGCATTCAGA 1260
OY 1261 AAGTTAATGAGTGTGTTTCCAGAAAGTGAAGCTGTAGTTCGTAGTCACTGACATGATG 1320
Db 1261 AAGTTAATGAGTGTGTTTCCAGAAAGTGAAGCTGTAGTTCGTAGTCACTGACATGATG 1320
OY 1321 GGGAGTCTGAATCAAAATGCCAAAGTAGTGTATGAGAGTCTTAAATGAGTAGATG 1380
Db 1321 GGGAGTCTGAATCAAAATGCCAAAGTAGTGTATGAGAGTCTTAAATGAGTAGATG 1380
OY 1381 AATATCTGCTCTCAGAGAAATAGACTTACTGGCCAGTATGCTCCTGAGGGCTTAA 1440
Db 1381 AATATCTGCTCTCAGAGAAATAGACTTACTGGCCAGTATGCTCCTGAGGGCTTAA 1440
OY 1441 TATGTAAAAGTGAAGAGTTCACTCCAAATCAGTAGAGTAATAATTGAAGCAAAATAT 1500
Db 1441 TATGTAAAAGTGAAGAGTTCACTCCAAATCAGTAGAGTAATAATTGAAGCAAAATAT 1500
OY 1501 TTGGGAAAACCTATGGGAAGGCAAGGCTCCCACTTAAGCCATGTAAGTGAATG 1560
Db 1501 TTGGGAAAACCTATGGGAAGGCAAGGCTCCCACTTAAGCCATGTAAGTGAATG 1560
OY 1561 TAATTAATGAGCATTTGTTTACTGAGCCAGATPATACAGAGGCTCCCTCACAAATA 1620
Db 1561 TAATTAATGAGCATTTGTTTACTGAGCCAGATPATACAGAGGCTCCCTCACAAATA 1620
OY 1621 AATTAAGCGTAAAAGAGACCTACATCAGGCTTCATCTCTGAGATTTTATCAAGAAAG 1680
Db 1621 AATTAAGCGTAAAAGAGACCTACATCAGGCTTCATCTCTGAGATTTTATCAAGAAAG 1680
OY 1681 CAGATTTGGCAGTCTAAAAGACTCCTGAAATGATTAATCAGGAACTAAACCAAGGAGC 1740
Db 1681 CAGATTTGGCAGTCTAAAAGACTCCTGAAATGATTAATCAGGAACTAAACCAAGGAGC 1740
OY 1741 AGAATGCTCAAGTGAATATTACTAATAGTGTATGATGAGATTAACCAAAAGGTGAT 1800
Db 1741 AGAATGCTCAAGTGAATATTACTAATAGTGTATGATGAGATTAACCAAAAGGTGAT 1800
OY 1801 CTATTCAGATGAGAAAAATCTTAACCCAAATAGAAATCACTCGAAAAAGATCTGTTTCA 1860
Db 1801 CTATTCAGATGAGAAAAATCTTAACCCAAATAGAAATCACTCGAAAAAGATCTGTTTCA 1860
OY 1861 AAACGAAAGCTGAACCTATTAAGCAGAGTATTAAGCAATATGGAACCTCGAATTAATTC 1920
Db 1861 AAACGAAAGCTGAACCTATTAAGCAGAGTATTAAGCAATATGGAACCTCGAATTAATTC 1920
OY 1921 ACAATTCAAAAGCAGCTTAATAAAGATAGGCTGAGAGGAAGTCTTCTACAGGCAATTC 1980
Db 1921 ACAATTCAAAAGCAGCTTAATAAAGATAGGCTGAGAGGAAGTCTTCTACAGGCAATTC 1980
OY 1981 ATGCGCTTGAACCTAGTACGTAGAAATCTTAAGCCACCTAATTGTACTGAATTCGAAA 2040
Db 1981 ATGCGCTTGAACCTAGTACGTAGAAATCTTAAGCCACCTAATTGTACTGAATTCGAAA 2040
OY 2041 TTGATATGTTGTTCTAGCAGTGAAGATTAAGAAAAAAAGTACCAACCAATGCGCAGTCA 2100
Db 2041 TTGATATGTTGTTCTAGCAGTGAAGATTAAGAAAAAAAGTACCAACCAATGCGCAGTCA 2100
OY 2101 GGCACAGCAGAACTACAACTCATGGAAGTAAAGAACCTGCACTGAGGCAAGAAAGA 2160
Db 2101 GGCACAGCAGAACTACAACTCATGGAAGTAAAGAACCTGCACTGAGGCAAGAAAGA 2160
OY 2161 GTTAAACGCAAAATGAGACAGACAGTAAAGACATGACAGTATCTTTCCAGAGCTGA 2220
Db 2161 GTTAAACGCAAAATGAGACAGACAGTAAAGACATGACAGTATCTTTCCAGAGCTGA 2220
OY 2221 AGTTAAACAATGACACTGCTGTTCTTTACTAAGTGTCAAAATACAGTGAATTAAGAAAT 2280
Db 2221 AGTTAAACAATGACACTGCTGTTCTTTACTAAGTGTCAAAATACAGTGAATTAAGAAAT 2280
OY 2281 TTGTCAATCTAGCCTTCCAGAGAAAGAAAGAGAACTAGAAACAGTTAAAGTGT 2340
Db 2281 TTGTCAATCTAGCCTTCCAGAGAAAGAAAGAGAACTAGAAACAGTTAAAGTGT 2340
OY 2341 CTAAATTAAGTGAAGACCCCAAGATCTCATGTTAAGTGAAGAAAGGTTTTGCAAACTG 2400
Db 2341 CTAAATTAAGTGAAGACCCCAAGATCTCATGTTAAGTGAAGAAAGGTTTTGCAAACTG 2400
OY 2401 AAAAGATCTGTAGAGAGTACATATTTCTTTGGTACTGTGACTGATTTATGACACTCAG 2460
Db 2401 AAAAGATCTGTAGAGAGTACATATTTCTTTGGTACTGTGACTGATTTATGACACTCAG 2460
OY 2461 AAAGATCTGCTTACTGGAAGTTAGCACCTCTGGAAGAGCAAAACAGAACTAAAT 2520
Db 2461 AAAGATCTGCTTACTGGAAGTTAGCACCTCTGGAAGAGCAAAACAGAACTAAAT 2520
OY 2521 GTGTAGTCAAGTGTGAGCATTTTGAAGAACCCCAAGGACTAATTCATGTTTCCAAAG 2580
Db 2521 GTGTAGTCAAGTGTGAGCATTTTGAAGAACCCCAAGGACTAATTCATGTTTCCAAAG 2580
OY 2581 ATTAATGAATGACACAGAGGCTTTAATATCATGTTGGACATGAAAGTTAACACAGTC 2640
Db 2581 ATTAATGAATGACACAGAGGCTTTAATATCATGTTGGACATGAAAGTTAACACAGTC 2640
OY 2641 GGGAAACAAGCATAGAAATGGAAGAAAGTGAAGTCTGATCTCAGATTTGCAAGATACAT 2700
Db 2641 GGGAAACAAGCATAGAAATGGAAGAAAGTGAAGTCTGATCTCAGATTTGCAAGATACAT 2700
OY 2701 TCAAGGTTTCAAAGGCCAGCATTTGCTCTGTTTCAAATCCAGAAATGCAAGAGAG 2760
Db 2701 TCAAGGTTTCAAAGGCCAGCATTTGCTCTGTTTCAAATCCAGAAATGCAAGAGAG 2760
OY 2761 AATGTCCAACTTCTCTGCCCCTCTGCGGCTCTTAAAGAAACCAAGTCCAAAGTCACT 2820
Db 2761 AATGTCCAACTTCTCTGCCCCTCTGCGGCTCTTAAAGAAACCAAGTCCAAAGTCACT 2820
OY 2821 TTGAATGTGAACAAAAGAAAGAAATCAAGAAAGAAAGTATGATTAATCAAGCCGTAC 2880
Db 2821 TTGAATGTGAACAAAAGAAAGAAATCAAGAAAGAAAGTATGATTAATCAAGCCGTAC 2880
OY 2881 AGACAGTTAATATCACTGAGGCTTCTGTGTGTGTCAGAAAGATTAAGCAGTTGATA 2940
Db 2881 AGACAGTTAATATCACTGAGGCTTCTGTGTGTGTCAGAAAGATTAAGCAGTTGATA 2940
OY 2941 ATGCCAAATGATGATCAAGAGAGGCTGATGTTTGTATATCTCAGTTCAGAGGCA 3000
Db 2941 ATGCCAAATGATGATCAAGAGAGGCTGATGTTTGTATATCTCAGTTCAGAGGCA 3000
OY 3001 ACGAAACTGAGCTCAATTACTCCAAATTAACATGAGACTTTTACAAAACCCATATCGATAC 3060
Db 3001 ACGAAACTGAGCTCAATTACTCCAAATTAACATGAGACTTTTACAAAACCCATATCGATAC 3060
OY 3061 CACCACCTTTTCCATCAAGTCAATTTGTTAAACTAAAGTAAAGAAAAATCTGCTGAGAG 3120
Db 3061 CACCACCTTTTCCATCAAGTCAATTTGTTAAACTAAAGTAAAGAAAAATCTGCTGAGAG 3120
OY 3121 AAAACTTTGAGAGACATTTCAATGTCACTGGAAGAGAAATGGAATGGAACATTCCAA 3180
Db 3121 AAAACTTTGAGAGACATTTCAATGTCACTGGAAGAGAAATGGAATGGAACATTCCAA 3180
OY 3181 GTACAGTGAAGCAATTAAGCTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3240
Db 3181 GTACAGTGAAGCAATTAAGCTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3240
OY 3241 CAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3300
Db 3241 CAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3300
OY 3301 TAGGTTCCAGTATGAGAAACATTTCAAGCAGAACTAGTGAAGAAAGAGGCCCCAAATGCA 3360
Db 3301 TAGGTTCCAGTATGAGAAACATTTCAAGCAGAACTAGTGAAGAAAGAGGCCCCAAATGCA 3360
```

```
OY 3361 ATGCTATGCTTAGATTAGGGGTTTTGCAACCTGAGGCTTATAAACAAGCTTCTCGGAA 3420
|||||
DB 3361 ATGCTATGCTTAGATTAGGGGTTTTGCAACCTGAGGCTTATAAACAAGCTTCTCGGAA 3420
OY 3421 GTAATTGTAGAGTCTCTGAATAAATAAAGCAAGATATAGAGAGTGTTCAGACTGTAA 3480
|||||
DB 3421 GTAATTGTAGAGTCTCTGAATAAATAAAGCAAGATATAGAGAGTGTTCAGACTGTAA 3480
OY 3481 ATACAGATTTCTCTCATATCTGATTTGATTAAGTACAGAGCTATGGGAGATGTC 3540
|||||
DB 3481 ATACAGATTTCTCTCATATCTGATTTGATTAAGTACAGAGCTATGGGAGATGTC 3540
OY 3541 ATGATCTCAGGTTTGTCTGAGACACCTGATGACCTGTAGATGATGTGAATAAAG 3600
|||||
DB 3541 ATGATCTCAGGTTTGTCTGAGACACCTGATGACCTGTAGATGATGTGAATAAAG 3600
OY 3601 AAGATACGATTTCTGTAATAATGACATTAAGGAAGTTCTGCTGTTTACCAAAAGCG 3660
|||||
DB 3601 AAGATACGATTTCTGTAATAATGACATTAAGGAAGTTCTGCTGTTTACCAAAAGCG 3660
OY 3661 TCCAGAAAGAGAGCTTAGCAGAGAGTCCAGGCCCTTTCACCATACATTTGGCTCAG 3720
|||||
DB 3661 TCCAGAAAGAGAGCTTAGCAGAGAGTCCAGGCCCTTTCACCATACATTTGGCTCAG 3720
OY 3661 TCCAGAAAGAGAGCTTAGCAGAGAGTCCAGGCCCTTTCACCATACATTTGGCTCAG 3720
|||||
DB 3661 TCCAGAAAGAGAGCTTAGCAGAGAGTCCAGGCCCTTTCACCATACATTTGGCTCAG 3720
OY 3721 GTTACCGAAGAGGGGCCAAGAAATTAGAGTCTCAGAGAGACTTATCTAGTGAGATG 3780
|||||
DB 3721 GTTACCGAAGAGGGGCCAAGAAATTAGAGTCTCAGAGAGACTTATCTAGTGAGATG 3780
OY 3781 AAGAGCTTCCCTGCTTCCACACTGTTATTTGGTAAGTAAACAATATACCTTCCAGT 3840
|||||
DB 3781 AAGAGCTTCCCTGCTTCCACACTGTTATTTGGTAAGTAAACAATATACCTTCCAGT 3840
OY 3841 CTACTAGGCATAGCACCGTGTGTACCGAGTGTCTGTAGAAACACAGAGAGAAATTAT 3900
|||||
DB 3841 CTACTAGGCATAGCACCGTGTGTACCGAGTGTCTGTAGAAACACAGAGAGAAATTAT 3900
OY 3901 TATATTGTAAGATATAGCTTAATATGCTGAGTACAGAGTAATATTGGCAAGGCAATCTC 3960
|||||
DB 3901 TATATTGTAAGATATAGCTTAATATGCTGAGTACAGAGTAATATTGGCAAGGCAATCTC 3960
OY 3961 AGGAACATCACCTTAGTAGAGGAAACAATGTTGCTGAGTGTGTTTCTTACAGTGA 4020
|||||
DB 3961 AGGAACATCACCTTAGTAGAGGAAACAATGTTGCTGAGTGTGTTTCTTACAGTGA 4020
OY 4021 GTGAATTGGAAGACTGCACTCAAAATACAAACACCCAGAGTCCCTTCTTGATTGCTCT 4080
|||||
DB 4021 GTGAATTGGAAGACTGCACTCAAAATACAAACACCCAGAGTCCCTTCTTGATTGCTCT 4080
OY 4081 CCAACAACATAGAGCATCAGTGTGAAGCCAGAGGAGTGTGTGAGTGACAGAAATTGG 4140
|||||
DB 4081 CCAACAACATAGAGCATCAGTGTGAAGCCAGAGGAGTGTGTGAGTGACAGAAATTGG 4140
OY 4141 TTTGAGATGATGAGAAAGAGAGAGCGGCTTGAAGAAATATATGAAGAGCAAGACA 4200
|||||
DB 4141 TTTGAGATGATGAGAAAGAGAGAGCGGCTTGAAGAAATATATGAAGAGCAAGACA 4200
OY 4201 TGGATTCAAACTTAGTGAGAGCAGCATCTGGGTGTGAGAGTGAACAAGCGCTCTGAAG 4260
|||||
DB 4201 TGGATTCAAACTTAGTGAGAGCAGCATCTGGGTGTGAGAGTGAACAAGCGCTCTGAAG 4260
OY 4261 ACTGCTCAGGAGCTATCTCTCAGAGTGAAGTATTTTACCACTCAGAGAGGATACCATGC 4320
|||||
DB 4261 ACTGCTCAGGAGCTATCTCTCAGAGTGAAGTATTTTACCACTCAGAGAGGATACCATGC 4320
OY 4321 AACATTAACCTGATTAAGGCTCCAGCAGAAATGGCTGAAGTGAAGCTGTGTAAGACAGC 4380
|||||
DB 4321 AACATTAACCTGATTAAGGCTCCAGCAGAAATGGCTGAAGTGAAGCTGTGTAAGACAGC 4380
OY 4381 ATGGAGCGACGCTTCTTAACAGCTAACCTTCATCATTAAGTACTCTTCTGCCCTTGGAG 4440
|||||
DB 4381 ATGGAGCGACGCTTCTTAACAGCTAACCTTCATCATTAAGTACTCTTCTGCCCTTGGAG 4440
OY 4441 ACCTGCGAATCCAGAACAAAGCACAATCAGAAAAAGCAGTATTAACCTTCACAGAAAAAGTA 4500
|||||
```

Db 5521 GCACAGGTGTCACACCAATTTGCTGTGTGACGACCATGCTGACAGACAGACAAATGGCT 5580  
 Oy 5581 TCCATGCAATTTGGCAGATGTGTGAGCACCCTGTGTGACCCGAGAGTGGTGTGACA 5640  
 Db 5581 TCCATGCAATTTGGCAGATGTGTGAGCACCCTGTGTGACCCGAGAGTGGTGTGACA 5640  
 Oy 5641 GTGAGCACTTACCAGTGGCAGAGCTGACACCTTACCTGATACCCAGATCCCCACA 5700  
 Db 5641 GTGAGCACTTACCAGTGGCAGAGCTGACACCTTACCTGATACCCAGATCCCCACA 5700  
 Oy 5701 GCCACTACTGA 5711  
 Db 5701 GCCACTACTGA 5711  
 RESULT 4  
 AAT17495  
 ID AAT17495 standard; cDNA: 5914 BP.  
 XX AAT17495;  
 AC AAT17495;  
 XX 02-OCT-1996 (first entry)  
 DT Mutated BRCA1 coding sequence from PM09.  
 XX  
 DE Cancer therapy; breast and ovarian cancer predisposing gene; immunogen;  
 KW antibody production; germline alteration; probe; lesion neoplasia; human;  
 KW gene therapy; protein replacement therapy; protein mimetic; BRCA1; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 120..5711  
 FT /\*tag= a  
 FT /product= BRCA1 mutant  
 FT 2201  
 FT mutation  
 FT /\*tag= b  
 FT /note= "C to T mutation"  
 XX  
 PN WO9605306-A2.  
 XX  
 PD 22-FEB-1996.  
 XX  
 PE 11-AUG-1995; 95WO-US10202.  
 XX  
 PR 07-JUN-1995; 95US-0483553.  
 PR 12-AUG-1994; 94US-0289221.  
 PR 02-SEP-1994; 94US-0300266.  
 PR 16-SEP-1994; 94US-0308104.  
 PR 29-NOV-1994; 94US-0348824.  
 PR 24-MAR-1995; 95US-0409305.  
 PR 07-JUN-1995; 95US-0480784.  
 XX  
 PA (MIRI-) MYRIAD GENETICS INC.  
 PA (CANC-) CANCER INST.  
 PA (RECH-) CENT RECH DU CHUL.  
 XX  
 PI Durocher F, Emi M, Nakamura Y, Simard J, Shattuck-Eidens DK.  
 DR WPI; 1996-139702/14.  
 DR P-PSDB; AAR81481.  
 XX  
 XX New nucleic acid and polypeptide for mutant or polymorphic BRCA1  
 PT gene - for diagnosis and therapy of human breast and ovarian cancer  
 PT and for diagnosing pre-disposition to these cancers  
 XX  
 PS Claim 1; 218pp; English.  
 XX  
 CC AAT17439-T17453 and AAT17455-T17529 represent mutations of the human  
 CC breast and ovarian cancer predisposing gene (BRCA1) (for wild type see  
 CC AAT17438). Proteins encoded by these mutations (see AAR81483-R81487 and  
 CC AAR81489-R81546) can be used as immunogens for antibody production.  
 CC These mutant BRCA1 genes have at least 1 mutation or polymorphism in

CC comparison to the wild type sequence. By detecting a germline alteration  
 CC in the wild type BRCA1 gene, a predisposition for breast and ovarian  
 CC cancer can be diagnosed. In one method, BRCA1 mRNA isolated from a tissue  
 CC sample from a subject has a probe, corresponding to a fragment of the  
 CC wild type sequence (or an allele-specific probe for one of these  
 CC mutations), added to it. The conditions allow for hybridisation of the  
 CC probe to the mRNA, and any hybridisation which occurs is detected.  
 CC Alternatively the BRCA1 gene in the tissue sample is isolated, and a  
 CC shift in electrophoretic mobility of single stranded DNA from the sample  
 CC on a non-denaturing polyacrylamide gel indicates a mutation. These  
 CC methods of detection can also diagnose a lesion neoplasia associated with  
 CC the BRCA1 locus. The methods may be used in gene therapy, protein  
 CC replacement therapy and protein mimetics, and may be used to screen for  
 CC drugs in cancer therapy.  
 XX  
 SQ Sequence 5914 BP: 2006 A; 1155 C; 1316 G; 1437 T; 0 other:  
 Query Match 100.0%; Score 5709.4; DB 17; Length 5914;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 5710; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Oy 1 AGCTGCGTGAACATCTTCTGACACCCCGACAGGCTGTGGGTTCTCAGATACTGGCC 60  
 Db 1 AGCTGCGTGAACATCTTCTGACACCCCGACAGGCTGTGGGTTCTCAGATACTGGCC 60  
 Oy 61 CCTGCGCTCAGAGAGCCCTTACACCTCTGCTGGGTAAAGTTCAATGGAACAGAAAGAA 120  
 Db 61 CCTGCGCTCAGAGAGCCCTTACACCTCTGCTGGGTAAAGTTCAATGGAACAGAAAGAA 120  
 Oy 121 TGGATTATCTGCTCTTCCGCTTGAAGAGTACAAATGTCTAATGCTTATGAGAAA 180  
 Db 121 TGGATTATCTGCTCTTCCGCTTGAAGAGTACAAATGTCTAATGCTTATGAGAAA 180  
 Oy 181 TCTTAGAGTGTCCATCTGTGTGAGTGTATCAAGAAAGTCTCTCCAAAGTGTGACC 240  
 Db 181 TCTTAGAGTGTCCATCTGTGTGAGTGTATCAAGAAAGTCTCTCCAAAGTGTGACC 240  
 Oy 241 ACATATTTTGGCAATTTTGGCAATGCTGAAACCTTCTCAACACAGAAAGGGCTTACACGT 300  
 Db 241 ACATATTTTGGCAATTTTGGCAATGCTGAAACCTTCTCAACACAGAAAGGGCTTACACGT 300  
 Oy 301 GTCCCTTATGTAAAGATGATATTAACCAAAAGAGCTTACAGAAAGTATGAGATTAGTGC 360  
 Db 301 GTCCCTTATGTAAAGATGATATTAACCAAAAGAGCTTACAGAAAGTATGAGATTAGTGC 360  
 Oy 361 AACTTGTGGAAGACCTATGAAATCATTTGTGCTTTTCAAGCTTGACACAGGTTGGAGT 420  
 Db 361 AACTTGTGGAAGACCTATGAAATCATTTGTGCTTTTCAAGCTTGACACAGGTTGGAGT 420  
 Oy 421 ATGCAACAGCTATTAATTTTGCAGAAAAGAAATTAATCTCCGACACATCTAAAGATG 480  
 Db 421 ATGCAACAGCTATTAATTTTGCAGAAAAGAAATTAATCTCCGACACATCTAAAGATG 480  
 Oy 481 AAGTTCTATCATCCAAAGTATGGGCTACAGAAACCGTGCAGAAAGCTTCTACAGAGTG 540  
 Db 481 AAGTTCTATCATCCAAAGTATGGGCTACAGAAACCGTGCAGAAAGCTTCTACAGAGTG 540  
 Oy 541 AACCCGAAATCCTTCTCTGAGAGAAACCACTCTCAGTGTCTCAACTCTTAACCTTGAA 600  
 Db 541 AACCCGAAATCCTTCTCTGAGAGAAACCACTCTCAGTGTCTCAACTCTTAACCTTGAA 600  
 Oy 601 CTGTGAGAACTCTGAGAGCAAGAGGAGATACCACTCAAAAGACGCTCTCTACAAATG 660  
 Db 601 CTGTGAGAACTCTGAGAGCAAGAGGAGATACCACTCAAAAGACGCTCTCTACAAATG 660  
 Oy 661 AATTGGATCTGATTTCTTGAAGATACCGTTAATAGGCAACTTATGAGAGTGGAG 720  
 Db 661 AATTGGATCTGATTTCTTGAAGATACCGTTAATAGGCAACTTATGAGAGTGGAG 720  
 Oy 721 ATCAAGATTTGTTCAATATCAACCCCTCAAGAGACAGGATGAATAGTTGGATTGCTG 780  
 Db 721 ATCAAGATTTGTTCAATATCAACCCCTCAAGAGACAGGATGAATAGTTGGATTGCTG 780

QY 781 CAAAAAGGCTGCTGTGTAATTTCTGAGAGGATGTACAAATCTGAACATCATCAAC 840  
DB 781 CAAAAAGGCTGCTGTGTAATTTCTGAGAGGATGTACAAATCTGAACATCATCAAC 840  
QY 841 CCGATATATATGATTTGACACACCTGAGAGGCTGACGCTGAGAGGCTATCCAGAAAAGT 900  
DB 841 CCGATATATATGATTTGACACACCTGAGAGGCTGACGCTGAGAGGCTATCCAGAAAAGT 900  
QY 901 ATCAGGATAGTCTGTTTCAAACTTGCATGTGAGGACATGGGACAAATCTCATGCA 960  
DB 901 ATCAGGATAGTCTGTTTCAAACTTGCATGTGAGGACATGGGACAAATCTCATGCA 960  
QY 961 GCTCATATACAGCATGAGACAGCATTTATTACTCATAAAGACAGATGATGTAGAAA 1020  
DB 961 GCTCATATACAGCATGAGACAGCATTTATTACTCATAAAGACAGATGATGTAGAAA 1020  
QY 1021 AGGCTGAATTCGTGAATTAAGCAACAGGCTGGCTTAGCAAGAGGCAACATCAAGAT 1080  
DB 1021 AGGCTGAATTCGTGAATTAAGCAACAGGCTGGCTTAGCAAGAGGCAACATCAAGAT 1080  
QY 1081 GGGCTGGAAGTAAAGAAACATGTATATGATAGGCGGACTCCAGCAGCAAAAAAGTAG 1140  
DB 1081 GGGCTGGAAGTAAAGAAACATGTATATGATAGGCGGACTCCAGCAGCAAAAAAGTAG 1140  
QY 1141 ATCTGAATGCTGATCCCTGTGTGAGAGAAAGATGGAATAGCAGAAACTGCCATGCT 1200  
DB 1141 ATCTGAATGCTGATCCCTGTGTGAGAGAAAGATGGAATAGCAGAAACTGCCATGCT 1200  
QY 1201 CAGAGATCCTAGAGATCTGGAAGATGTCTCTTGATTAACACTTAATAGCAGCTTGA 1260  
DB 1201 CAGAGATCCTAGAGATCTGGAAGATGTCTCTTGATTAACACTTAATAGCAGCTTGA 1260  
QY 1261 AAGTTAATGAGTGTTCAGAGAGTATGATGATGATGATGATGATGATGATGATGATG 1320  
DB 1261 AAGTTAATGAGTGTTCAGAGAGTATGATGATGATGATGATGATGATGATGATGATG 1320  
QY 1321 GGGAGTCTGAATTAATGCCAAAGTACGTATGATGATGATGATGATGATGATGATG 1380  
DB 1321 GGGAGTCTGAATTAATGCCAAAGTACGTATGATGATGATGATGATGATGATGATG 1380  
QY 1381 AATATCTGCTTCTTCAAGAAATAGACTTACTGGCAGATGATGATGATGATGATGAT 1440  
DB 1381 AATATCTGCTTCTTCAAGAAATAGACTTACTGGCAGATGATGATGATGATGATGAT 1440  
QY 1441 TATGTAAGAGTGAAGAGTTCACCTCAAAATCAGTAGAGTAATATTGAAGACAAAATAT 1500  
DB 1441 TATGTAAGAGTGAAGAGTTCACCTCAAAATCAGTAGAGTAATATTGAAGACAAAATAT 1500  
QY 1501 TTGGGAAAACCTATCGGAGAGGCAAGCCTCCCAACTTAAGCATGTAACTGAATATC 1560  
DB 1501 TTGGGAAAACCTATCGGAGAGGCAAGCCTCCCAACTTAAGCATGTAACTGAATATC 1560  
QY 1561 TATATATAGGAGATTTGTTACTGAGGCAAGATTAACAAAGGCGTCCCTCACAATA 1620  
DB 1561 TATATATAGGAGATTTGTTACTGAGGCAAGATTAACAAAGGCGTCCCTCACAATA 1620  
QY 1621 AATTAAGGCTAAAGAGACTACATCAGGCTTCATCTGAGAGATTTTATCAAGAAG 1680  
DB 1621 AATTAAGGCTAAAGAGACTACATCAGGCTTCATCTGAGAGATTTTATCAAGAAG 1680  
QY 1681 CAGATTTGGCAGTTCAAAAGACTCTGAAATGATTAATCAAGGAACTAACCAAGGAGC 1740  
DB 1681 CAGATTTGGCAGTTCAAAAGACTCTGAAATGATTAATCAAGGAACTAACCAAGGAGC 1740  
QY 1741 AGAATGCTCAAGTGAATATTAATGATGATGATGATGATGATGATGATGATGATGAT 1800  
DB 1741 AGAATGCTCAAGTGAATATTAATGATGATGATGATGATGATGATGATGATGATGAT 1800  
QY 1801 CTATTCAGATGAGAAATCTTAACCAATAGAAATCAGCAAAAGATCTGCTTCA 1860  
DB 1801 CTATTCAGATGAGAAATCTTAACCAATAGAAATCAGCAAAAGATCTGCTTCA 1860  
QY 1861 AAACGAAAGCTGAACCTATTAAGCAGCATATTAAGCAATATGAATCTGAATTAATATCC 1920  
DB 1861 AAACGAAAGCTGAACCTATTAAGCAGCATATTAAGCAATATGAATCTGAATTAATATCC 1920  
QY 1921 ACAATTTCAAAAGCACCCTAAAAAGATATGCTGAGAGGAAAGTCTTACCAGCATATTC 1980  
DB 1921 ACAATTTCAAAAGCACCCTAAAAAGATATGCTGAGAGGAAAGTCTTACCAGCATATTC 1980  
QY 1981 ATGCGCTTGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040  
DB 1981 ATGCGCTTGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040  
QY 2041 TTGATAGTGTCTTACAGTGAAGATTAAGAAAAAAGTAAACCAATGCGAGTCA 2100  
DB 2041 TTGATAGTGTCTTACAGTGAAGATTAAGAAAAAAGTAAACCAATGCGAGTCA 2100  
QY 2101 GGCACACGAAACCTTACACTCATGAGAGTAAAGAACTGTGAACTGAGGCAAGAGA 2160  
DB 2101 GGCACACGAAACCTTACACTCATGAGAGTAAAGAACTGTGAACTGAGGCAAGAGA 2160  
QY 2161 GTAAACAGCCAAATGAACAGACAAAGTAAAGACATGACAGTATCTTCCAGAGCTGA 2220  
DB 2161 GTAAACAGCCAAATGAACAGACAAAGTAAAGACATGACAGTATCTTCCAGAGCTGA 2220  
QY 2221 AGTTAACAAATGCACTGCTTCTTACTAGATTTCAATACAGTGAACCTTAAGAAAT 2280  
DB 2221 AGTTAACAAATGCACTGCTTCTTACTAGATTTCAATACAGTGAACCTTAAGAAAT 2280  
QY 2281 TTGTCATCTTACCTTCCAGAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGTAAAGT 2340  
DB 2281 TTGTCATCTTACCTTCCAGAGAGAAAGAAAGAAAGAAAGAAAGAAAGTAAAGT 2340  
QY 2341 CTATATATGCTGAAGACCCCAAGATCTCATGTTAAGTGGAGAAAGGTTTGCACACG 2400  
DB 2341 CTATATATGCTGAAGACCCCAAGATCTCATGTTAAGTGGAGAAAGGTTTGCACACG 2400  
QY 2401 AAGATCTGTAGAGAGTACGATTTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2460  
DB 2401 AAGATCTGTAGAGAGTACGATTTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2460  
QY 2461 AAGATCTGTAGAGAGTACGATTTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2520  
DB 2461 AAGATCTGTAGAGAGTACGATTTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2520  
QY 2521 GTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2580  
DB 2521 GTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2580  
QY 2581 ATATATGAAATGACACAGAGGCTTTAAGTATCATTGGAGCATGAGTTAACACAGTC 2640  
DB 2581 ATATATGAAATGACACAGAGGCTTTAAGTATCATTGGAGCATGAGTTAACACAGTC 2640  
QY 2641 GGGAAACAGCATGAAATGAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2700  
DB 2641 GGGAAACAGCATGAAATGAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2700  
QY 2701 TCAAGGTTTCAAGCCCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760  
DB 2701 TCAAGGTTTCAAGCCCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760  
QY 2761 AATGTGCAACATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820  
DB 2761 AATGTGCAACATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820  
QY 2821 TTGAATGCTCAACAAAGGAAAGAAATCAAGGAAAGTGAAGTGAAGTGAAGTGAAGTGA 2880  
DB 2821 TTGAATGCTCAACAAAGGAAAGAAATCAAGGAAAGTGAAGTGAAGTGAAGTGAAGTGA 2880  
QY 2881 AGACAGTTATATACATGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2940  
DB 2881 AGACAGTTATATACATGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2940  
QY 2941 ATGCCAAATGATATCAAGGAGGCTGAGTTTGTCTATCATCTGACATTCAGAGGCA 3000  
DB 2941 ATGCCAAATGATATCAAGGAGGCTGAGTTTGTCTATCATCTGACATTCAGAGGCA 3000

D 2941 ATGCCAATGTAGATCAAGAGAGCGCTGAGTTTTGTCTATCATCTCAGTTCAGAGCA 3000  
Q ACAGAACTGAGCTCATCTACTCCAAATTAACATGAGACTTTTACAAACCCTATTCGTATAC 3060  
D ACACAACTGGAGCTCATCTACTCCAAATTAACATGAGACTTTTACAAACCCTATTCGTATAC 3060  
Q CACCACTTTTCCCATCACTCATCTTTGTAAACTAATGTAAGAAATCTGCTAGAG 3120  
D CACCACTTTTCCCATCACTCATCTTTGTAAACTAATGTAAGAAATCTGCTAGAG 3120  
Q AAAACTTTGAGAGCAATTCATCACTGACCTGAAAGAGAAATGGCAATGAGAACATTC 3180  
D AAAACTTTGAGAGCAATTCATCACTGACCTGAAAGAGAAATGGCAATGAGAACATTC 3180  
Q GTACAGTGAAGCAATTTAGCCGTAATTAACATTTAGAGAAATGTTTTTAAAGAACCCAGCT 3240  
D GTACAGTGAAGCAATTTAGCCGTAATTAACATTTAGAGAAATGTTTTTAAAGAACCCAGCT 3240  
Q CAAGCAATTTAATGAAGTGGTTCCAGTACTAATGAAGTGGCTCCAGTATTAATGAAA 3300  
D CAAGCAATTTAATGAAGTGGTTCCAGTACTAATGAAGTGGCTCCAGTATTAATGAAA 3300  
Q TAGGTCAGTGAATGAAGCAATTCAGAGCAAGTACTAGTAAACAGAGGCCCAAAATGGA 3360  
D TAGGTCAGTGAATGAAGCAATTCAGAGCAAGTACTAGTAAACAGAGGCCCAAAATGGA 3360  
Q ATGCATGCTAGTATTAAGGGGTTTGGCAACCTGAGGCTATTAACAAAGTCTTCTGGA 3420  
D ATGCATGCTAGTATTAAGGGGTTTGGCAACCTGAGGCTATTAACAAAGTCTTCTGGA 3420  
Q GTATATGTAGCATCCCGAAATTAAGAAAGCAAGAAATGAAGAGTATGTTTACAGCTTTA 3480  
D GTATATGTAGCATCCCGAAATTAAGAAAGCAAGAAATGAAGAGTATGTTTACAGCTTTA 3480  
Q ATACAGATTTCTCTCCATATCTGATTTCAATTAAGTAAAGAGTATGAGAGTATGAGT 3540  
D ATACAGATTTCTCTCCATATCTGATTTCAATTAAGTAAAGAGTATGAGAGTATGAGT 3540  
Q ATGCATGCTAGTATTAAGGGGTTTGGCAACCTGAGGCTATTAACAAAGTCTTCTGGA 3600  
D ATGCATGCTAGTATTAAGGGGTTTGGCAACCTGAGGCTATTAACAAAGTCTTCTGGA 3600  
Q AAGTACTAGTATTTGCGAAATGACATTAAGAAAGTCTGCTGTTTATAGCAAAAGCG 3660  
D AAGTACTAGTATTTGCGAAATGACATTAAGAAAGTCTGCTGTTTATAGCAAAAGCG 3660  
Q AAGTACTAGTATTTGCGAAATGACATTAAGAAAGTCTGCTGTTTATAGCAAAAGCG 3660  
D AAGTACTAGTATTTGCGAAATGACATTAAGAAAGTCTGCTGTTTATAGCAAAAGCG 3660  
Q TCCAGAAAGAGAGCTTAGCAGAGAGTCTAGCCCTTTACCCATACCATTTGGCTAGG 3720  
D TCCAGAAAGAGAGCTTAGCAGAGAGTCTAGCCCTTTACCCATACCATTTGGCTAGG 3720  
Q GTTACCGAAGAGGGGCCAAGAAATTAGAGTCTCAGAGAGAACTTATCTAGTAGAGATG 3780  
D GTTACCGAAGAGGGGCCAAGAAATTAGAGTCTCAGAGAGAACTTATCTAGTAGAGATG 3780  
Q AAGAGCTTCCCTGCTTCCAACTGTTATTTGGTAAAGTAAATATACCTTTCAGT 3840  
D AAGAGCTTCCCTGCTTCCAACTGTTATTTGGTAAAGTAAATATACCTTTCAGT 3840  
Q CTACTAGAGCTAGCAGCGTTGCTACCGAGTGTCTGTCTAAGAACACAGAGAGATTTAT 3900  
D CTACTAGAGCTAGCAGCGTTGCTACCGAGTGTCTGTCTAAGAACACAGAGAGATTTAT 3900  
Q TATCATTTGAAGATAGCTTAATGACTGAGTAAACAGAGTAAATTTGGCAAGGCAATCTC 3960  
D TATCATTTGAAGATAGCTTAATGACTGAGTAAACAGAGTAAATTTGGCAAGGCAATCTC 3960  
Q AGGAATCATCCTTTAGTGAAGAGAAACAATGTTCTGTAGCTTTTCTTCCACAGTGA 4020  
D AGGAATCATCCTTTAGTGAAGAGAAACAATGTTCTGTAGCTTTTCTTCCACAGTGA 4020  
Q GTGAATTTGAAGACTTGAGTGAAGTAAATGAGAACACCCAGATCTTCTTATTTGGTCTT 4080  
D GTGAATTTGAAGACTTGAGTGAAGTAAATGAGAACACCCAGATCTTCTTATTTGGTCTT 4080

Q CCAACAAATGAGCATCATCTGTAAGCCAGAGAGTTGGTCTAGTGAAGCAATTTGG 4140  
D CCAACAAATGAGCATCATCTGTAAGCCAGAGAGTTGGTCTAGTGAAGCAATTTGG 4140  
Q TTTGATGATGAAGAAAGAGAGAGCGGCTTGGAAAGAAATTAATCAAGAAAGAGCA 4200  
D TTTGATGATGAAGAAAGAGAGAGCGGCTTGGAAAGAAATTAATCAAGAAAGAGCA 4200  
Q TGGATTTAAACTTAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4260  
D TGGATTTAAACTTAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4260  
Q ACTGCTCAGGGCTATCTCTCAGAGTGAATTTTAACTTACAGAGAGAGAGAGAGAG 4320  
D ACTGCTCAGGGCTATCTCTCAGAGTGAATTTTAACTTACAGAGAGAGAGAGAGAG 4320  
Q AACATACCTGATTAAGCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4380  
D AACATACCTGATTAAGCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4380  
Q ATGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4440  
D ATGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4440  
Q ATGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4500  
D ATGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4500  
Q ACCTGCGAAATCCAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4560  
D ACCTGCGAAATCCAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4560  
Q GAGATAGTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4620  
D GAGATAGTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4620  
Q GAGATAGTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4680  
D GAGATAGTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4680  
Q GAGATAGTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4740  
D GAGATAGTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4740  
Q GAGATAGTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4800  
D GAGATAGTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4800  
Q GAGATAGTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4860  
D GAGATAGTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4860  
Q GAGATAGTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4920  
D GAGATAGTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4920  
Q GAGATAGTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4980  
D GAGATAGTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4980  
Q GAGATAGTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5040  
D GAGATAGTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5040  
Q GAGATAGTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5100  
D GAGATAGTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5100  
Q GAGATAGTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5160  
D GAGATAGTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5160



QY 5161 CTGAAGAGACTACTCATGTTGTTATGAAAAACAGATGCTGAGTTGTTGTTGAACGGACAC 5220  
 |||||  
 DB 5161 CTGAAGAGACTACTCATGTTGTTATGAAAAACAGATGCTGAGTTGTTGTTGAACGGACAC 5220  
 QY 5221 TGAATATTTTCTAGGAATTTGCCGGAGAAATGGTAGTACCTATTTTGGGTGACCC 5280  
 |||||  
 DB 5221 TGAATATTTTCTAGGAATTTGCCGGAGAAATGGTAGTACCTATTTTGGGTGACCC 5280  
 QY 5281 AGCTCTTAAAGAAAAAATGCTGAATGAGCATGATTTTGAAGTCAGAGGAGATGTGG 5340  
 |||||  
 DB 5281 AGCTCTTAAAGAAAAAATGCTGAATGAGCATGATTTTGAAGTCAGAGGAGATGTGG 5340  
 QY 5341 TCAATGGAAGAAACCAAGAGTCCAAAGGAGACAGAGATCCAGACAGAAAGATCT 5400  
 |||||  
 DB 5341 TCAATGGAAGAAACCAAGAGTCCAAAGGAGACAGAGATCCAGACAGAAAGATCT 5400  
 QY 5401 TCAGGGGGCTAGAAATCTGTTGTTATGAGGCCCTTTCACCAACATGCCCCAGATCACTG 5460  
 |||||  
 DB 5401 TCAGGGGGCTAGAAATCTGTTGTTATGAGGCCCTTTCACCAACATGCCCCAGATCACTG 5460  
 QY 5461 AATGATGTGACAGCTGTGGTCTTCTGTTGTTGAAGAGCTTTCATCATTCACCCCTTG 5520  
 |||||  
 DB 5461 AATGATGTGACAGCTGTGGTCTTCTGTTGTTGAAGAGCTTTCATCATTCACCCCTTG 5520  
 QY 5521 GCACAGGTGTCACCAATTTGTTGTTGTCAGCCAGATGCTTGCAGAGAGACAAATGGCT 5580  
 |||||  
 DB 5521 GCACAGGTGTCACCAATTTGTTGTTGTCAGCCAGATGCTTGCAGAGAGACAAATGGCT 5580  
 QY 5581 TCCATGCCATTTGGGCGAGATGTTGAGAGGCACCTGTGTGACCCGAGAGTGGTGTGGACA 5640  
 |||||  
 DB 5581 TCCATGCCATTTGGGCGAGATGTTGAGAGGCACCTGTGTGACCCGAGAGTGGTGTGGACA 5640  
 QY 5641 GTGTAGCACTTACAGTCCAGAGAGTGCAGACCTACCTGATACCCAGATCCCCACACA 5700  
 |||||  
 DB 5641 GTGTAGCACTTACAGTCCAGAGAGTGCAGACCTACCTGATACCCAGATCCCCACACA 5700  
 QY 5701 GCCACTACTGA 5711  
 |||||  
 DB 5701 GCCACTACTGA 5711  
 RESULT 5  
 AAV46463  
 ID AAV46463 standard; cDNA: 5711 BP.  
 AC AAV46463;  
 XX  
 DT 18-NOV-1998 (first entry)  
 XX  
 DE Human BRCA1 omi2 polymorphism #6 cDNA.  
 XX  
 KW BRCA1: omi2; human: breast and ovarian cancer predisposing gene;  
 KM polymorphism; susceptibility; anti-oncogene; tumour suppressor;  
 KW chromosome 17q; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH key location/Qualifiers  
 FT CDS 120..5711  
 FT /\*tag= a  
 FT /product= "BRCA1 omi2 protein"  
 FT variation 4427  
 FT /tag= b  
 FT /note= "This polymorphic variation can be a T or C  
 nucleotide"  
 XX  
 PN US5750400-A.  
 XX  
 PD 12-MAY-1998.  
 XX  
 PE 12-FEB-1997; 97US-0798691.  
 XX  
 PR 12-FEB-1996; 96US-0598591.

PR 12-FEB-1997; 97US-0798691.  
 XX  
 PA (ONCO-) ONCORMED INC.  
 XX  
 PI Allen AC, Alvares CP, Critz BS, Murphy PD, Olson SJ;  
 PI Scheller DB, Zeng B;  
 DR WP1: 1998-296774/26.  
 XX  
 PT BRCA1 omi2 gene coding sequences - useful for distinguishing between  
 PT polymorphisms and mutation(s) in the screening for disposition to  
 PT breast or ovarian cancer  
 XX  
 PS Claim 2e; Page -: 54pp; English.  
 CC This sequence encodes a human BRCA1 (breast and ovarian cancer  
 CC predisposing gene) omi2 gene in which a polymorphic variation occurs at  
 CC nucleotide 4427. This sequence and other polymorphic variations of this  
 CC sequence are useful for the identification of an individual who may or  
 CC may not have an increased susceptibility to breast or ovarian cancer.  
 CC The sequences used identify gene changes which are due to polymorphisms,  
 CC rather than cancer causing mutations. BRCA1 is an anti-oncogene (tumour  
 CC suppressor) which is involved in genetic inheritance of cancers,  
 CC especially breast and ovarian cancer. It is found at human chromosome  
 CC 17q which is known to be linked to cancer susceptibility, especially  
 CC breast cancer. Cells containing a mutation in this gene lose the  
 CC wild-type function of BRCA1 and are more susceptible to cancers.  
 CC NOTE: This sequence does not appear in the specification but has been  
 CC created from the wild type BRCA1 omi2 gene represented in AAV46449.  
 XX  
 S0 Sequence 5711 BP; 1956 A; 1098 C; 1274 G; 1382 T; 1 other;  
 Query Match 100.0%; Score 5709; DB 19; Length 5711;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 5709; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 AGCTGCGTGAAGCTTCCGTCGACCCGACACAGGCTGTGGGGTTTCTCAGATACCTGGGCC 60  
 |||||  
 DB 1 AGCTGCGTGAAGCTTCCGTCGACCCGACACAGGCTGTGGGGTTTCTCAGATACCTGGGCC 60  
 QY 61 CCTGGCGTCAAGAGGCTTCCACCTGCTGCTGGGTAAAGTTCAATGGAACAGAAAGAAA 120  
 |||||  
 DB 61 CCTGGCGTCAAGAGGCTTCCACCTGCTGCTGGGTAAAGTTCAATGGAACAGAAAGAAA 120  
 QY 121 TGGATTTATCTGCTCTTCCGCTTGAAGAGTCAAAATGTCATTAATCTTTCAGAGAAA 180  
 |||||  
 DB 121 TGGATTTATCTGCTCTTCCGCTTGAAGAGTCAAAATGTCATTAATCTTTCAGAGAAA 180  
 QY 181 TCTTAGAGTGTCCATCTGCTGAGTTGATCAAGGAACCTGTCTCCAAAGTGTGACC 240  
 |||||  
 DB 181 TCTTAGAGTGTCCATCTGCTGAGTTGATCAAGGAACCTGTCTCCAAAGTGTGACC 240  
 QY 241 ACATATTTTGGCAATTTTGCATGCTGAACCTTCTACCCAGAGAAAGGCTTTCACAGT 300  
 |||||  
 DB 241 ACATATTTTGGCAATTTTGCATGCTGAACCTTCTACCCAGAGAAAGGCTTTCACAGT 300  
 QY 301 GTCCCTTATGTAAGATGATATACCAAAAGAGGCTTCAAGAAAGTCCGAGATTTTGTGC 360  
 |||||  
 DB 301 GTCCCTTATGTAAGATGATATACCAAAAGAGGCTTCAAGAAAGTCCGAGATTTTGTGC 360  
 QY 361 AACCTGTGGAAGAGCTATTTGAAGAAATCATTTGCTTTACAGTTGACACAGATTTGGAGT 420  
 |||||  
 DB 361 AACCTGTGGAAGAGCTATTTGAAGAAATCATTTGCTTTACAGTTGACACAGATTTGGAGT 420  
 QY 421 ATGCAAAACAGCTATATTTTGGCAAAAAGAAATTAAGTCTGCAAGATCTTAAGATG 480  
 |||||  
 DB 421 ATGCAAAACAGCTATATTTTGGCAAAAAGAAATTAAGTCTGCAAGATCTTAAGATG 480  
 QY 481 AAGTTTCTATCATTCGAAGATATGGGCTACAGAAACCTGCCAAAGAACTCTACAGAGT 540  
 |||||  
 DB 481 AAGTTTCTATCATTCGAAGATATGGGCTACAGAAACCTGCCAAAGAACTCTACAGAGT 540  
 QY 541 AACCCGAAATTCCTTCTGTCAGGAACACAGTCTCAGTCTCAACTCTTAACCTTGAA 600



```
|||||
Db 541 AACCCGAAATCCCTCCCTCCAGAAACCACTCTCACTGTCACAACTCTAACCTTGAA 600
QY CTCTGGAACCTGAGAGCAACAGCGGATACAACTCAAAAGACTGCTGTACTTG 660
Db 601 CTCTGGAACCTGAGAGCAACAGCGGATACAACTCAAAAGACTGCTGTACTTG 660
QY 661 AATTGGATCTGATCTCTGAGATACCTTAATAGGCAACTTATTGAGTGGAG 720
Db 661 AATTGGATCTGATCTCTGAGATACCTTAATAGGCAACTTATTGAGTGGAG 720
QY 721 ATCAGAAATGTTACAATACCCCTCAGGAACAGGATGAATCAGTTGGATTCTG 780
Db 721 ATCAGAAATGTTACAATACCCCTCAGGAACAGGATGAATCAGTTGGATTCTG 780
QY 781 CAAAAAGGCTGCTGGAATTTCTGAGAGGATTAACAAATACGACATCATAC 840
Db 781 CAAAAAGGCTGCTGGAATTTCTGAGAGGATTAACAAATACGACATCATAC 840
QY 841 CCAATTAATTAATGATTTGAGACCACTGAGAGCTGAGAGGATCCAGAAAGT 900
Db 841 CCAATTAATTAATGATTTGAGACCACTGAGAGCTGAGAGGATCCAGAAAGT 900
QY 901 ATCAGGCTAGTCTCTGTTCAACTGTCATGTCGAGCCATGTCACAAATCTCATCCA 960
Db 901 ATCAGGCTAGTCTCTGTTCAACTGTCATGTCGAGCCATGTCACAAATCTCATCCA 960
QY 961 GCTCATTAACAGCATGAGACAGCGATTATTACTACATAAGACAGATTAATGATAAA 1020
Db 961 GCTCATTAACAGCATGAGACAGCGATTATTACTACATAAGACAGATTAATGATAAA 1020
QY 1021 AGGCTGAATCTCTAATTAAGCAACAGCGCTGCTAGCAAGAGGCAACATTAACGAT 1080
Db 1021 AGGCTGAATCTCTAATTAAGCAACAGCGCTGCTAGCAAGAGGCAACATTAACGAT 1080
QY 1081 GGGCTGGAATTAAGCAACATGTAATGATAGGCGACTCCACAGACAGAAAAAGGTAG 1140
Db 1081 GGGCTGGAATTAAGCAACATGTAATGATAGGCGACTCCACAGACAGAAAAAGGTAG 1140
QY 1141 ATCTGAATGCTGATCCCTGCTGAGAGAAAGATGAATTAAGACAAACTGCGATGCT 1200
Db 1141 ATCTGAATGCTGATCCCTGCTGAGAGAAAGATGAATTAAGACAAACTGCGATGCT 1200
QY 1201 CAGAGATCTCTAGAGATGTAAGATGTCCTTGATTAACATAATAGCAGATTGAGA 1260
Db 1201 CAGAGATCTCTAGAGATGTAAGATGTCCTTGATTAACATAATAGCAGATTGAGA 1260
QY 1261 AAGTTAATGATGCTTTTCCAGAAAGTATGATGCTGATGCTGATGATGATG 1320
Db 1261 AAGTTAATGATGCTTTTCCAGAAAGTATGATGCTGATGCTGATGATGATG 1320
QY 1321 GGGAGCTGGAATCAAAATGCCAAGTATGATGATGATGATGATGATGATGATG 1380
Db 1321 GGGAGCTGGAATCAAAATGCCAAGTATGATGATGATGATGATGATGATGATG 1380
QY 1381 AATATCTGCTCTCTAGAGAAATAGACTTACTGAGCAGTATGATGATGATGATG 1440
Db 1381 AATATCTGCTCTCTAGAGAAATAGACTTACTGAGCAGTATGATGATGATGATG 1440
QY 1441 TATGTAATGATGAAAGTCTCAATCAATCAGTAGAGATTAATATTGAAGACAAATAT 1500
Db 1441 TATGTAATGATGAAAGTCTCAATCAATCAGTAGAGATTAATATTGAAGACAAATAT 1500
QY 1501 TTGGGAAACCTTATCGGAAGAGCAAGCTCCCAACTTAAGCATGTAAGTGAATATC 1560
Db 1501 TTGGGAAACCTTATCGGAAGAGCAAGCTCCCAACTTAAGCATGTAAGTGAATATC 1560
QY 1561 TAATTAATGAGAGATTTGTTACTGAGCACAGATTAATCAAGAGCGCTCCCTCACAATA 1620
Db 1561 TAATTAATGAGAGATTTGTTACTGAGCACAGATTAATCAAGAGCGCTCCCTCACAATA 1620
QY 1621 AATTTAAAGCTAAAGAGACCTACATCAGGCTTACCTGAGATTTTATCAAGAAAG 1680
|||||
Db 1621 AATTTAAAGCTAAAGAGACCTACATCAGGCTTACCTGAGATTTTATCAAGAAAG 1680
QY 1681 CAGATTGGCAGTTCAAAAGACTCCTGAAATGATTAATACAGGAATTAACCAAGCGAGC 1740
Db 1681 CAGATTGGCAGTTCAAAAGACTCCTGAAATGATTAATACAGGAATTAACCAAGCGAGC 1740
QY 1741 AGAATGGTCAAGTGAATTAATTAATAGTGTCTATGAGAAATTAACAAAGGTGATTT 1800
Db 1741 AGAATGGTCAAGTGAATTAATTAATAGTGTCTATGAGAAATTAACAAAGGTGATTT 1800
QY 1801 CTATTCAGATGAGAAAAATCTTAACCAATAGATGATCAGTGAAGAAATCTGCTTTCA 1860
Db 1801 CTATTCAGATGAGAAAAATCTTAACCAATAGATGATCAGTGAAGAAATCTGCTTTCA 1860
QY 1861 AAACGAAGCTGAACCTATAGAGAGCTATAGCAATATGGAAGCTGGAATTAATATTC 1920
Db 1861 AAACGAAGCTGAACCTATAGAGAGCTATAGCAATATGGAAGCTGGAATTAATATTC 1920
QY 1921 ACAATTCAAAAAGCCTTAAGAAAGATAGGCTGAGAGAGGATCTTACAGAGCATATTC 1980
Db 1921 ACAATTCAAAAAGCCTTAAGAAAGATAGGCTGAGAGAGGATCTTACAGAGCATATTC 1980
QY 1981 ATGCGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2040
Db 1981 ATGCGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2040
QY 2041 TTGATGATGTTGTTAGAGAGTGAAGATTAAGAAAAAGTATACCAACCAATGGCAGTCA 2100
Db 2041 TTGATGATGTTGTTAGAGAGTGAAGATTAAGAAAAAGTATACCAACCAATGGCAGTCA 2100
QY 2101 GGCACAGCAGAAACCTCAACTCATGAGAGGTAAGAAACCTGCACTGAGACCAAGAGA 2160
Db 2101 GGCACAGCAGAAACCTCAACTCATGAGAGGTAAGAAACCTGCACTGAGACCAAGAGA 2160
QY 2161 GTAACAAAGCCAAATGAACAGCAAGTAAAGACATGACAGTGAATTTCCACAGCTGA 2220
Db 2161 GTAACAAAGCCAAATGAACAGCAAGTAAAGACATGACAGTGAATTTCCACAGCTGA 2220
QY 2221 AGTTAACAATGACACCTGCTGTTCTTTACTAAGTCTCAATATACAGGAATTTAAGAT 2280
Db 2221 AGTTAACAATGACACCTGCTGTTCTTTACTAAGTCTCAATATACAGGAATTTAAGAT 2280
QY 2281 TTGTCAATCTAGGCTTCCAGAGAGAAAAAGAGAAACCTAGAAACCACTTAAAGTGT 2340
Db 2281 TTGTCAATCTAGGCTTCCAGAGAGAAAAAGAGAAACCTAGAAACCACTTAAAGTGT 2340
QY 2341 CTATTAATGCTGAAGACCCCAAGATCTCATGTTAAGTGAAGAGGTTTGGCAACTG 2400
Db 2341 CTATTAATGCTGAAGACCCCAAGATCTCATGTTAAGTGAAGAGGTTTGGCAACTG 2400
QY 2401 AAAGATCTGTAAGAGTAGAGCAATTTCAATGATGATGATGATGATGATGATGATG 2460
Db 2401 AAAGATCTGTAAGAGTAGAGCAATTTCAATGATGATGATGATGATGATGATGATG 2460
QY 2461 AAAGATCTGTAAGAGTAGAGCAATTTCAATGATGATGATGATGATGATGATGATG 2520
Db 2461 AAAGATCTGTAAGAGTAGAGCAATTTCAATGATGATGATGATGATGATGATGATG 2520
QY 2521 GTGTGATGCTAGTGTGAGCAATTTAAAGCCCAAGGAGCTAATTCATGATGTTGCTGAAG 2580
Db 2521 GTGTGATGCTAGTGTGAGCAATTTAAAGCCCAAGGAGCTAATTCATGATGTTGCTGAAG 2580
QY 2581 ATATTAATGATGAGACAGAGAGGCTTTAAGTATGATGATGATGATGATGATGATGATG 2640
Db 2581 ATATTAATGATGAGACAGAGAGGCTTTAAGTATGATGATGATGATGATGATGATGATG 2640
QY 2641 GGGAAACCAAGCATTAAGAAATGGAAGAGTGAACCTGATGATGATGATGATGATGATG 2700
Db 2641 GGGAAACCAAGCATTAAGAAATGGAAGAGTGAACCTGATGATGATGATGATGATGATGATG 2700
QY 2701 TCAAGGTTTCAAGGCGGAGCTGATGCTGCTGTTTCAAAATCCAGGAAATGCAAGAGG 2760
Db 2701 TCAAGGTTTCAAGGCGGAGCTGATGCTGCTGTTTCAAAATCCAGGAAATGCAAGAGG 2760
```

OY 2761 AATGTGCAACATCTCTGCCACTCTGGGCTCTTAAAGAAACAAGTCCAAAAGTCACCTT 2820  
|||||  
Db 2761 AATGTGCAACATCTCTGCCACTCTGGGCTCTTAAAGAAACAAGTCCAAAAGTCACCTT 2820  
OY 2821 TTGAATGTGAACAAAGGAAGAAATCAGAGAAAGATGAGTCTAATATCAAGCTGTAC 2880  
|||||  
Db 2821 TTGAATGTGAACAAAGGAAGAAATCAGAGAAAGATGAGTCTAATATCAAGCTGTAC 2880  
OY 2881 AGACAGTTAATATCACTGCAGGCTTCTGTGGTGTGTCAGAAAGATAGCCAGTTGATA 2940  
|||||  
Db 2881 AGACAGTTAATATCACTGCAGGCTTCTGTGGTGTGTCAGAAAGATAGCCAGTTGATA 2940  
OY 2941 ATGCCAAATGTAGTATCAAGAGAGGCTAGGTTTGTCTATCATCTCACTTCAGAGCA 3000  
|||||  
Db 2941 ATGCCAAATGTAGTATCAAGAGAGGCTAGGTTTGTCTATCATCTCACTTCAGAGCA 3000  
OY 3001 AGGAAACTGACATCTACTCCAAATTAACATGAGCTTTTACAAAACCCATATCTATAC 3060  
|||||  
Db 3001 AGGAAACTGACATCTACTCCAAATTAACATGAGCTTTTACAAAACCCATATCTATAC 3060  
OY 3061 CACCACTTTTCCCATCAAGTCAATTTGTTAAACTAAATGTAAAGAAAATCTGCTAGAG 3120  
|||||  
Db 3061 CACCACTTTTCCCATCAAGTCAATTTGTTAAACTAAATGTAAAGAAAATCTGCTAGAG 3120  
OY 3121 AAAACTTTGAGGAACATTCATATGTCACCTGAAAGAGAAATGGAAATGAGAACATTTCCA 3180  
|||||  
Db 3121 AAAACTTTGAGGAACATTCATATGTCACCTGAAAGAGAAATGGAAATGAGAACATTTCCA 3180  
OY 3181 GTACAGTGAACACATTAAGCCGTAATTAACCTTAGAGAAATGTTTTTAAAGAGCCAGCT 3240  
|||||  
Db 3181 GTACAGTGAACACATTAAGCCGTAATTAACCTTAGAGAAATGTTTTTAAAGAGCCAGCT 3240  
OY 3241 CAAGCAATATTAATGAAGTAGGTTCCAGTACTAATGAAGTGGGCTCCAGTATTAATGAA 3300  
|||||  
Db 3241 CAAGCAATATTAATGAAGTAGGTTCCAGTACTAATGAAGTGGGCTCCAGTATTAATGAA 3300  
OY 3301 TAGGTTCCAGTATGAAAAATTCAGAGCAAGACTAGTAGAAACAGAGGCCAAATTTGA 3360  
|||||  
Db 3301 TAGGTTCCAGTATGAAAAATTCAGAGCAAGACTAGTAGAAACAGAGGCCAAATTTGA 3360  
OY 3361 ATGCTATGCTTAGATTAGAGGGTTTTGCAACCTGAGGCTCTATTAACAAGCTTCTGAGAA 3420  
|||||  
Db 3361 ATGCTATGCTTAGATTAGAGGGTTTTGCAACCTGAGGCTCTATTAACAAGCTTCTGAGAA 3420  
OY 3421 GTAATTTGTAAGCATCCTGAAATTAAGAAACAAGATATGAGAAGTAGTTCAGACGTCTA 3480  
|||||  
Db 3421 GTAATTTGTAAGCATCCTGAAATTAAGAAACAAGATATGAGAAGTAGTTCAGACGTCTA 3480  
OY 3481 ATACAGATTTCTCTCCATATCTGATTTTCAAGATTAAGTAAAGCTATGAGAGTATGATGC 3540  
|||||  
Db 3481 ATACAGATTTCTCTCCATATCTGATTTTCAAGATTAAGTAAAGCTATGAGAGTATGATGC 3540  
OY 3541 ATGCAATTCAGGTTTGTGAGACACGCTGATGAGTGTAGATGATGATGATGATGATGATG 3600  
|||||  
Db 3541 ATGCAATTCAGGTTTGTGAGACACGCTGATGAGTGTAGATGATGATGATGATGATGATG 3600  
OY 3601 AAGATACATAGTTTGTGTAAGATTAAGAAAGTCTGCTGTTTGTAGCAAAAGCG 3660  
|||||  
Db 3601 AAGATACATAGTTTGTGTAAGATTAAGAAAGTCTGCTGTTTGTAGCAAAAGCG 3660  
OY 3661 TTCCAGAAAGAGAGCTTAGCAGAGAGTCTAGCCCTTTCACCCATACACATTTGGCTCAG 3720  
|||||  
Db 3661 TTCCAGAAAGAGAGCTTAGCAGAGAGTCTAGCCCTTTCACCCATACACATTTGGCTCAG 3720  
OY 3721 GTTACGGAAGAGGGGCGCAGAAATTAAGATTCCTCAAGAGAACTATCTAGAGAGATG 3780  
|||||  
Db 3721 GTTACGGAAGAGGGGCGCAGAAATTAAGATTCCTCAAGAGAACTATCTAGAGAGATG 3780  
OY 3781 AAGAGCTTCCCTGCTTCCAAACTTGTATTGGTAAAGTAACATTAATACCTTCTCAGT 3840  
|||||  
Db 3781 AAGAGCTTCCCTGCTTCCAAACTTGTATTGGTAAAGTAACATTAATACCTTCTCAGT 3840

OY 3841 CTACTAGCATATACACCGCTTGTACCGAGTGTCTGTCTAAGAAACACAGAGAGAAATTAAT 3900  
|||||  
Db 3841 CTACTAGCATATACACCGCTTGTACCGAGTGTCTGTCTAAGAAACACAGAGAGAAATTAAT 3900  
OY 3901 TATCATTTGAAGATFAGCTTAATAGACTGCAGTAAACAGATTAATATTGGCAAGGATCTC 3960  
|||||  
Db 3901 TATCATTTGAAGATFAGCTTAATAGACTGCAGTAAACAGATTAATATTGGCAAGGATCTC 3960  
OY 3961 AGCAATCATCTAGTGAAGAAACAAGAAATGTTCTCTAGCTGTGTTTCTTACAGTGA 4020  
|||||  
Db 3961 AGCAATCATCTAGTGAAGAAACAAGAAATGTTCTCTAGCTGTGTTTCTTACAGTGA 4020  
OY 4021 GTGAATTTGAAGACTTGACTGCAATTAACAACCCAGGATCTCTTGTGATTTGCTCT 4080  
|||||  
Db 4021 GTGAATTTGAAGACTTGACTGCAATTAACAACCCAGGATCTCTTGTGATTTGCTCT 4080  
OY 4081 CCAAAACAATGAGGCAATGCTGAAAGCAAGGAGTGTGCTGAGTGAAGAAATGG 4140  
|||||  
Db 4081 CCAAAACAATGAGGCAATGCTGAAAGCAAGGAGTGTGCTGAGTGAAGAAATGG 4140  
OY 4141 TTTTCAGATGATGAAGAAAGAGAAAGCGGCTTGGAAAGAAATTAATCAAGAGCAAGCA 4200  
|||||  
Db 4141 TTTTCAGATGATGAAGAAAGAGAAAGCGGCTTGGAAAGAAATTAATCAAGAGCAAGCA 4200  
OY 4201 TGGATTTCAAATCTTAGTGAAGCAGCATCTGGGTGTGAGAGTGAACAGCTCTCTGAAG 4260  
|||||  
Db 4201 TGGATTTCAAATCTTAGTGAAGCAGCATCTGGGTGTGAGAGTGAACAGCTCTCTGAAG 4260  
OY 4261 ACTGCTCAGGCTATCTCTCAAGTGAATTTTAAACCTCAAGCAAGAGGATACCATGC 4320  
|||||  
Db 4261 ACTGCTCAGGCTATCTCTCAAGTGAATTTTAAACCTCAAGCAAGAGGATACCATGC 4320  
OY 4321 AACATTAACCTGATTAAGAGCTCCAGCAGAAATGGCTAACTAGAAGCTGTGTAGAAGC 4380  
|||||  
Db 4321 AACATTAACCTGATTAAGAGCTCCAGCAGAAATGGCTAACTAGAAGCTGTGTAGAAGC 4380  
OY 4381 ATGGAGACGAGCTTCTTAACAGCTACCTTCCATTAAGTGAAGTCTCTGCTTGAAG 4440  
|||||  
Db 4381 ATGGAGACGAGCTTCTTAACAGCTACCTTCCATTAAGTGAAGTCTCTGCTTGAAG 4440  
OY 4441 ACCTGCGAAATTCAGAAACAAGCAGCATCAAGAAACAGTATTAATTCAGAGAAAGTA 4500  
|||||  
Db 4441 ACCTGCGAAATTCAGAAACAAGCAGCATCAAGAAACAGTATTAATTCAGAGAAAGTA 4500  
OY 4501 GTGAATACCTTATTAACCCAGAAATCCAGAAAGGCTTCTGCTGCAAGTTTGAAGTCTG 4560  
|||||  
Db 4501 GTGAATACCTTATTAACCCAGAAATCCAGAAAGGCTTCTGCTGCAAGTTTGAAGTCTG 4560  
OY 4561 CAGATAGTTCTTACAGTAAATAAGAAACAGAGAGTGAAGAGTATCCCTTCTAAT 4620  
|||||  
Db 4561 CAGATAGTTCTTACAGTAAATAAGAAACAGAGAGTGAAGAGTATCCCTTCTAAT 4620  
OY 4621 GCCCATCATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4680  
|||||  
Db 4621 GCCCATCATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4680  
OY 4681 ACTACCATCTCAAGAGAGCTATTAAGGTTTGTATGTGAGAGAGCAACAGCTGGAAG 4740  
|||||  
Db 4681 ACTACCATCTCAAGAGAGCTATTAAGGTTTGTATGTGAGAGAGCAACAGCTGGAAG 4740  
OY 4741 AGTCTGGGCGCACAGATTTGACGGAACATCTTACTTGTGCAAGGCAAGATCTAGAGGAA 4800  
|||||  
Db 4741 AGTCTGGGCGCACAGATTTGACGGAACATCTTACTTGTGCAAGGCAAGATCTAGAGGAA 4800  
OY 4801 CCCCTTACCGGAATGTGAATCAGCTCTCTGATGATGATGATGATGATGATGATGATG 4860  
|||||  
Db 4801 CCCCTTACCGGAATGTGAATCAGCTCTCTGATGATGATGATGATGATGATGATGATG 4860  
OY 4861 AAGACAGAGCCCGACAGATCTGCTGTGGAACATCAATCTTCAACCTCTGCAATTTGA 4920  
|||||  
Db 4861 AAGACAGAGCCCGACAGATCTGCTGTGGAACATCAATCTTCAACCTCTGCAATTTGA 4920  
OY 4921 AAGTTCGCCAATTTGAAAGTTGCAAGATCTGCCAGAGTCCAGCTGCTCATACTACTG 4980  
|||||

```
QY 2521 GTGTAGTCACTGTGCAGCATTTGAAAAACCCCAAGGAGCTAATTCATGTTGTTCCAAG 2580
D 2521 GTGTAGTCACTGTGCAGCATTTGAAAAACCCCAAGGAGCTAATTCATGTTGTTCCAAG 2580
QY 2581 ATATACAAATGACAGAAAGGCTTTAAGTATCCATTGGGACATGAAGTTACCACAGTC 2640
D 2581 ATATACAAATGACAGAAAGGCTTTAAGTATCCATTGGGACATGAAGTTACCACAGTC 2640
QY 2641 GGGAAACAGCATGTGAATGAGAAAGTGAACCTTGATGCTAGTATTTGACAAATCAT 2700
D 2641 GGGAAACAGCATGTGAATGAGAAAGTGAACCTTGATGCTAGTATTTGACAAATCAT 2700
QY 2701 TCAAGGTTTCAAGCGCCAGTCATTTGCTGTTTTCGAATCCAGAAATGACAGAAAGG 2760
D 2701 TCAAGGTTTCAAGCGCCAGTCATTTGCTGTTTTCGAATCCAGAAATGACAGAAAGG 2760
QY 2761 AATGTGCAACATTTCTGCGCCACTGCTGGGCTTTAAAGAAACAAAGTCCAAAGTCACTT 2820
D 2761 AATGTGCAACATTTCTGCGCCACTGCTGGGCTTTAAAGAAACAAAGTCCAAAGTCACTT 2820
QY 2821 TTGAATGTGAACAAAGAGAAATCAAGAAAGTGAAGTCAATATCAAGGCTGTAC 2880
D 2821 TTGAATGTGAACAAAGAGAAATCAAGAAAGTGAAGTCAATATCAAGGCTGTAC 2880
QY 2881 AGACAGTTAATATCATGTCAGAGGCTTTCCTGTGTGTGTCAGAAAGATAGCCAGTTGATA 2940
D 2881 AGACAGTTAATATCATGTCAGAGGCTTTCCTGTGTGTGTCAGAAAGATAGCCAGTTGATA 2940
QY 2941 ATGCGCAATGTAGTATCAAAAGAGGCTGTAGGTTTGTCTATCTCTCAGTTCCAGAGCA 3000
D 2941 ATGCGCAATGTAGTATCAAAAGAGGCTGTAGGTTTGTCTATCTCTCAGTTCCAGAGCA 3000
QY 3001 ACGAACTGTGACATCTACTCCAAATTAACATGAGCTTTTACAAACCCATATGCTATAC 3060
D 3001 ACGAACTGTGACATCTACTCCAAATTAACATGAGCTTTTACAAACCCATATGCTATAC 3060
QY 3061 CACCACATTTTCCCATCAATCAATGTTGTAAACATAATGTAAAGAAATCTGCTAGAG 3120
D 3061 CACCACATTTTCCCATCAATCAATGTTGTAAACATAATGTAAAGAAATCTGCTAGAG 3120
QY 3121 AAAAATTTGAGAAACATTCATGCTCAGCTGAAAGAGAAATGGGAAATGAGAAACATTCGAA 3180
D 3121 AAAAATTTGAGAAACATTCATGCTCAGCTGAAAGAGAAATGGGAAATGAGAAACATTCGAA 3180
QY 3181 GTACAGTGAGACAAATTAAGCCGTAATACATTAGAGAAATGTTTTAAAGAACCCAGCT 3240
D 3181 GTACAGTGAGACAAATTAAGCCGTAATACATTAGAGAAATGTTTTAAAGAACCCAGCT 3240
QY 3241 CAAGCAATATTAATGAAGTGTCCAGTACTATGAAGTGGGCTCCAGTATTAATGAAA 3300
D 3241 CAAGCAATATTAATGAAGTGTCCAGTACTATGAAGTGGGCTCCAGTATTAATGAAA 3300
QY 3301 TAGGTTCCAGTATGAAGAAATTCAGACAGAACTAGTAGAAACAGAGGCGCAAAATTTGA 3360
D 3301 TAGGTTCCAGTATGAAGAAATTCAGACAGAACTAGTAGAAACAGAGGCGCAAAATTTGA 3360
QY 3361 ATGCTATGCTTAGATTTAGGGGTTTTGCAACCTGAGGCTATTAACAAGATCTTCTGGA 3420
D 3361 ATGCTATGCTTAGATTTAGGGGTTTTGCAACCTGAGGCTATTAACAAGATCTTCTGGA 3420
QY 3421 GTAATTTGAAGATCTCTGAATTAAGAAAGCAAGATATGAAGAAAGTATGAGTCACTTTA 3480
D 3421 GTAATTTGAAGATCTCTGAATTAAGAAAGCAAGATATGAAGAAAGTATGAGTCACTTTA 3480
QY 3481 ATACAGATTTCTCTCCATATCTGATTTCAATTAACATTAGAACAGCCATAGGAAATGATC 3540
D 3481 ATACAGATTTCTCTCCATATCTGATTTCAATTAACATTAGAACAGCCATAGGAAATGATC 3540
QY 3541 ATGCAATCTAGGTTTGTGTGAGACACCTGATGACCTGTAGATGATGATGATGATGATGATG 3600
D 3541 ATGCAATCTAGGTTTGTGTGAGACACCTGATGACCTGTAGATGATGATGATGATGATGATG 3600
QY 3601 AAGATCTAGTGTGTGTAAGTATGACATTAAGAAAGTCTGCTGTTTTAGCAAAAGCG 3660
D 3601 AAGATCTAGTGTGTGTAAGTATGACATTAAGAAAGTCTGCTGTTTTAGCAAAAGCG 3660

D 3601 AAGATCTAGTGTGTGTAAGTATGACATTAAGAAAGTCTGCTGTTTTAGCAAAAGCG 3660
QY 3661 TCCAGAAAGGAGACTTAGCAGAGAGTCTTACCCCTTTCACCAATACATTTGGCTCAGG 3720
D 3661 TCCAGAAAGGAGACTTAGCAGAGAGTCTTACCCCTTTCACCAATACATTTGGCTCAGG 3720
QY 3721 GTTACCGAAGAGGGGCCAAGAAATTAAGTCTCTGAGAAAGAACCTTATCTAGTGAATG 3780
D 3721 GTTACCGAAGAGGGGCCAAGAAATTAAGTCTCTGAGAAAGAACCTTATCTAGTGAATG 3780
QY 3781 AAGAGCTTCCCTGCTTCCAAACACTTTGTTATTTGGTAAAGTAAACATATACCTTCACT 3840
D 3781 AAGAGCTTCCCTGCTTCCAAACACTTTGTTATTTGGTAAAGTAAACATATACCTTCACT 3840
QY 3841 CTACTAGGATAGCAGCCGTTGCTACCGAGTGTCTGTCTAGAACACAGAGAGAAATTTAT 3900
D 3841 CTACTAGGATAGCAGCCGTTGCTACCGAGTGTCTGTCTAGAACACAGAGAGAAATTTAT 3900
QY 3901 TATCATTTGAAGAAATAGCTTAATGACTGCAAGTACCAAGTAAATTTGCAAAAGCATCTC 3960
D 3901 TATCATTTGAAGAAATAGCTTAATGACTGCAAGTACCAAGTAAATTTGCAAAAGCATCTC 3960
QY 3961 AGAATCATCACCTTAGTGAAGAAACAAATGTTCTGCTAGCTGTTTCTTCCAGTGA 4020
D 3961 AGAATCATCACCTTAGTGAAGAAACAAATGTTCTGCTAGCTGTTTCTTCCAGTGA 4020
QY 4021 GTGAAATTTGGAAGACTTTACGTCAAATACAAACACCCAGGATCCCTTTCTGTTGCTT 4080
D 4021 GTGAAATTTGGAAGACTTTACGTCAAATACAAACACCCAGGATCCCTTTCTGTTGCTT 4080
QY 4081 CCAAAACAAATGAGGACATCTAGTGAAGAGGAGGAGTGGCTGAGGACAGAAAGATTTG 4140
D 4081 CCAAAACAAATGAGGACATCTAGTGAAGAGGAGGAGTGGCTGAGGACAGAAAGATTTG 4140
QY 4141 TTTTCAGATGATGAAGAAAGAGAAAGGAGGCTTGAAGAAATTAATCAAGAAAGCAAGCA 4200
D 4141 TTTTCAGATGATGAAGAAAGAGAAAGGAGGCTTGAAGAAATTAATCAAGAAAGCAAGCA 4200
QY 4201 TGGATTTAAACTTTAGGTGAGGACAGCATCTGGGCTGAGAGTGAAGAAACGCTCTGTAG 4260
D 4201 TGGATTTAAACTTTAGGTGAGGACAGCATCTGGGCTGAGAGTGAAGAAACGCTCTGTAG 4260
QY 4261 ACTGCTAGGAGCTTCCCTCTCAGAGTGAACATTTAAACACAGACAGAGGATTCATGAC 4320
D 4261 ACTGCTAGGAGCTTCCCTCTCAGAGTGAACATTTAAACACAGACAGAGGATTCATGAC 4320
QY 4321 AACATACTGTATTAAGCTCCAGCAGAAATGGCTGAAGTGAAGAGTGTATTAGAACAGC 4380
D 4321 AACATACTGTATTAAGCTCCAGCAGAAATGGCTGAAGTGAAGAGTGTATTAGAACAGC 4380
QY 4381 ATGGAGACCGACGCTTCTTAACAGCTACCTTTCATATAGTACTCTTCTGCTTGAAG 4440
D 4381 ATGGAGACCGACGCTTCTTAACAGCTACCTTTCATATAGTACTCTTCTGCTTGAAG 4440
QY 4441 ACCGTGGAATCCAGAACAAAGCAATCAGAAAGAGCAATTAACCTTCACAGAAAGTGA 4500
D 4441 ACCGTGGAATCCAGAACAAAGCAATCAGAAAGAGCAATTAACCTTCACAGAAAGTGA 4500
QY 4501 GTGAATACCTTATTAAGCCAGAAATCCAGAAAGGCTTTTCTGCTGACAACTTTGAGTGTCTG 4560
D 4501 GTGAATACCTTATTAAGCCAGAAATCCAGAAAGGCTTTTCTGCTGACAACTTTGAGTGTCTG 4560
QY 4561 CAGATAGTCTTACAGATTAAGAAATTAAGAACACAGAGAGTGAAGAGTCTTCCCTTTCTAAT 4620
D 4561 CAGATAGTCTTACAGATTAAGAAATTAAGAACACAGAGAGTGAAGAGTCTTCCCTTTCTAAT 4620
QY 4621 GCCCATCATTAAGATGATGAGGAGTACATGACAGTCTCTGAGAGCTTCACAAATGAA 4680
D 4621 GCCCATCATTAAGATGATGAGGAGTACATGACAGTCTCTGAGAGCTTCACAAATGAA 4680
QY 4681 ACTAACCATCTCAAGAGAGCTCATTAAGGTTGTGATGTGAGAGCAACAGCTGAG 4740
D 4681 ACTAACCATCTCAAGAGAGCTCATTAAGGTTGTGATGTGAGAGCAACAGCTGAG 4740
```

```

Db 4681 ACTACCATCTCAAGAGAGCCTATTAGGTGTGATGTGAGAGCAACAGCTGGAG 4740
OY 4741 AGTCTGGCCACACGATTTTGACGGAACATCTTACTTGCCAGAGCATCTAGAGGAA 4800
Db 4741 AGTCTGGCCACACGATTTTGACGGAACATCTTACTTGCCAGAGCATCTAGAGGAA 4800
OY 4801 CCCCTTACCTGGAATCTGGAATCAGCCCTTCTCTGTGATGACCTGAACTGATCTTCTG 4860
Db 4801 CCCCTTACCTGGAATCTGGAATCAGCCCTTCTCTGTGATGACCTGAACTGATCTTCTG 4860
OY 4861 AAGACAGAGCCCCAGAGTGCAGCTCGTGTGGCAACATACCATCTTCAACCTCTGCATGA 4920
Db 4861 AAGACAGAGCCCCAGAGTGCAGCTCGTGTGGCAACATACCATCTTCAACCTCTGCATGA 4920
OY 4921 AAGTTCGCCAATGAAGTTCAGAAATCTGCCAGATCCAGCTGCTCATACTACTG 4980
Db 4921 AAGTTCGCCAATGAAGTTCAGAAATCTGCCAGATCCAGCTGCTCATACTACTG 4980
OY 4981 ATACTGCTGGGTATTAATGCAATGGAAGAGTGTGAGCAGGAGAGAACAGAAATTGACAG 5040
Db 4981 ATACTGCTGGGTATTAATGCAATGGAAGAGTGTGAGCAGGAGAGAACAGAAATTGACAG 5040
OY 5041 CTTCAACAGAAAGGTCACAAAGAAATGTCATGGTGTGTGCTGCTTACCCAGAG 5100
Db 5041 CTTCAACAGAAAGGTCACAAAGAAATGTCATGGTGTGTGCTGCTTACCCAGAG 5100
OY 5101 AATTATGCTGCTGATCAAGTTCAGAAACACCATCAGTCTTACTTAATTA 5160
Db 5101 AATTATGCTGCTGATCAAGTTCAGAAACACCATCAGTCTTACTTAATTA 5160
OY 5161 CTGAGAGACTACTCATGTGTGTTATGAAACAGATGCTGAGTTGTGTGTAACGAGAC 5220
Db 5161 CTGAGAGACTACTCATGTGTGTTATGAAACAGATGCTGAGTTGTGTGTAACGAGAC 5220
OY 5221 TGAATATTTTCTAGGAATTTGGGAGGAAATGGGAGTACTTCTTGGTGACCC 5280
Db 5221 TGAATATTTTCTAGGAATTTGGGAGGAAATGGGAGTACTTCTTGGTGACCC 5280
OY 5281 AGTCTATTAAAGAAAGAAATGCTGAATGAGCATATTTTGAAGTCAGAGAGATGTG 5340
Db 5281 AGTCTATTAAAGAAAGAAATGCTGAATGAGCATATTTTGAAGTCAGAGAGATGTG 5340
OY 5341 TCAATGGAAGAAACCCCAAGGTCACAAAGCAGACAGAGATCCAGAGAGAAAGATCT 5400
Db 5341 TCAATGGAAGAAACCCCAAGGTCACAAAGCAGACAGAGATCCAGAGAGAAAGATCT 5400
OY 5401 TCAGGGGGCTAGAAATCTGTGCTATGGGCCCTTCAACAACATGCCACAGATCACTGG 5460
Db 5401 TCAGGGGGCTAGAAATCTGTGCTATGGGCCCTTCAACAACATGCCACAGATCACTGG 5460
OY 5461 AATGATGTTACAGCTGTGTGTGCTGTGTGTAAGAGAGCTTTCATCTTACCTTGG 5520
Db 5461 AATGATGTTACAGCTGTGTGTGCTGTGTGTAAGAGAGCTTTCATCTTACCTTGG 5520
OY 5521 GCACAGGTTCACACCAATTTGTTGTGACAGCCAGATGCTTGCAGAGAGACATGGCT 5580
Db 5521 GCACAGGTTCACACCAATTTGTTGTGACAGCCAGATGCTTGCAGAGAGACATGGCT 5580
OY 5581 TCCATGCAATTTGGGAGATGTGAGAGACCTGTGTGACCCAGAGTGGTGTGGACA 5640
Db 5581 TCCATGCAATTTGGGAGATGTGAGAGACCTGTGTGACCCAGAGTGGTGTGGACA 5640
OY 5641 GTGTAGCACTTACACAGTCCAGAGAGCTGGACACTACCTGATACCCAGATCCCCACA 5700
Db 5641 GTGTAGCACTTACACAGTCCAGAGAGCTGGACACTACCTGATACCCAGATCCCCACA 5700
OY 5701 GCCACTACTGA 5711
Db 5701 GCCACTACTGA 5711

```

RESULT 7  
AAV46458

```

ID AAV46458 standard; cDNA: 5711 BP.
XX
AC AAV46458:
XX
DT 18-NOV-1998 (first entry)
XX
DE Human BRCA1 omi2 polymorphism #1 cDNA.
XX
KW BRCA1; omi2; human; breast and ovarian cancer predisposing gene;
KW polymorphism; susceptibility; anti-oncogene; tumour suppressor;
KW chromosome 17q; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 120..5711
FT FT /*tag= a
FT FT /product= "BRCA1 omi2 protein"
FT FT variation
FT FT 2201
FT FT /*tag= b
FT FT /note= "this polymorphic variation can be a C or T
nucleotide"
XX
XX US5750400-A.
XX
XX 12-MAY-1998.
XX
XX 12-FEB-1997; 9705-0798691.
XX
XX 12-FEB-1996; 9605-0598591.
XX
XX 12-FEB-1997; 9705-0798691.
XX
XX (ONCO-) ONCORMED INC.
XX
PI Allen AC, Alvares CP, Critz BS, Murphy PD, Olson SJ;
PI Scheller DB, Zeng B;
XX
DR WP1: 1998-296774/26.
XX
PT BRCA1 omi2 gene coding sequences - useful for distinguishing between
PT polymorphisms and mutation(s) in the screening for disposition to
PT breast or ovarian cancer
XX
XX
XX Claim 2e: Page -: 54pp: English.
XX
CC This sequence encodes a human BRCA1 (breast and ovarian cancer
CC predisposing gene) omi2 gene in which a polymorphic variation occurs at
CC nucleotide 2201. This sequence and other polymorphic variations of this
CC sequence are useful for the identification of an individual who may or
CC may not have an increased susceptibility to breast or ovarian cancer.
CC The sequences used identify gene changes which are due to polymorphisms,
CC rather than cancer causing mutations. BRCA1 is an anti-oncogene (tumour
CC suppressor) which is involved in genetic inheritance of cancers,
CC especially breast and ovarian cancer. It is found at human chromosome
CC 17q which is known to be linked to cancer susceptibility, especially
CC breast cancer. Cells containing a mutation in this gene lose the
CC wild-type function of BRCA1 and are more susceptible to cancers.
CC NOTE: This sequence does not appear in the specification but has been
CC created from the wild type BRCA1 omi2 gene represented in AAV46449.
XX
XX
SQ Sequence 5711 BP; 1956 A; 1098 C; 1274 G; 1382 T; 1 other:
XX
Query Match 100.0%; Score 5709; DB 19; Length 5711;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5709; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 1 AGCTGCTGAGACTTCTCTGACCCGACACGAGCTGTGGGTTTCTCAGATAACTGGCC 60
Db 1 AGCTGCTGAGACTTCTCTGACCCGACACGAGCTGTGGGTTTCTCAGATAACTGGCC 60
OY 61 CCTGGCCTCAGAGAGGCTTACCCCTTCTGTGCTGTGGTAAAGTTCATTGGAACAGAAAGAA 120
Db 61 CCTGGCCTCAGAGAGGCTTACCCCTTCTGTGCTGTGGTAAAGTTCATTGGAACAGAAAGAA 120

```

121 TGAATTTATCTGCTTCGCGCTTGAGAGATACAAATGTCATTATGCTATGAGAAAA 180  
121 TGAATTTATCTGCTTCGCGCTTGAGAGATACAAATGTCATTATGCTATGAGAAAA 180  
181 TCTTAGAGTGTCCATCTGCTGAGGTTGATCAAGAACTGTCTCCCAAAAGTGTACC 240  
181 TCTTAGAGTGTCCATCTGCTGAGGTTGATCAAGAACTGTCTCCCAAAAGTGTACC 240  
241 ACATATTTTGCATTTTGCATGCTGTAACCTTCAACCCAGAGAAAGGCGCTCACAGT 300  
241 ACATATTTTGCATTTTGCATGCTGTAACCTTCAACCCAGAGAAAGGCGCTCACAGT 300  
301 GTCCCTTATGTAAGATGATTAACCAAAAGAGCCTTACAAAGAAAGTACAGATTTAGTC 360  
301 GTCCCTTATGTAAGATGATTAACCAAAAGAGCCTTACAAAGAAAGTACAGATTTAGTC 360  
361 AACTGTTGAAGAGCTTATGAAATCATTTGTGCTTTTACGCTTGACACAGGTTGAGT 420  
361 AACTGTTGAAGAGCTTATGAAATCATTTGTGCTTTTACGCTTGACACAGGTTGAGT 420  
421 ATGCAACAGCTATATTTTGCATTTTGCATGCTGTAACCTTCAACCCAGAGAAAGT 480  
421 ATGCAACAGCTATATTTTGCATTTTGCATGCTGTAACCTTCAACCCAGAGAAAGT 480  
481 AAGTTTCTATCATCAAAAGTATGAGGCTACAGAAACCGTCCCAAAAGCTTCTACAGAGT 540  
481 AAGTTTCTATCATCAAAAGTATGAGGCTACAGAAACCGTCCCAAAAGCTTCTACAGAGT 540  
541 AACCCGAAATCTCTCTCTGAGAGAAACAGCTGCTCAACTCTCAACCTTGGA 600  
541 AACCCGAAATCTCTCTCTGAGAGAAACAGCTGCTCAACTCTCAACCTTGGA 600  
601 CTGAGAGACCTGAGAGCAAGAGGAGATACCAACCTCAAAAGCCTCTCTACATG 660  
601 CTGAGAGACCTGAGAGCAAGAGGAGATACCAACCTCAAAAGCCTCTCTACATG 660  
661 AATTGGATCTGATTTCTCTGAGAGATACCGTTAATAGGCAACTTATGCGATGGAG 720  
661 AATTGGATCTGATTTCTCTGAGAGATACCGTTAATAGGCAACTTATGCGATGGAG 720  
721 ATCAAGATTTTCAAAATCAACCCCTCAAGAGACCGAGGATGAATCACTTGAATCTG 780  
721 ATCAAGATTTTCAAAATCAACCCCTCAAGAGACCGAGGATGAATCACTTGAATCTG 780  
781 CAAAAAGGCTGCTGTAATTTCTGAGAGGATGAATCACTTGAATCACTTGAATCTG 840  
781 CAAAAAGGCTGCTGTAATTTCTGAGAGGATGAATCACTTGAATCACTTGAATCTG 840  
841 CCAAGTAATATGATTTTGAACACCACTGAGAGCGCTGAGAGGATCCAGAAAAAGT 900  
841 CCAAGTAATATGATTTTGAACACCACTGAGAGCGCTGAGAGGATCCAGAAAAAGT 900  
901 ATCAAGGATGCTGCTTCAAACTTGCATGAGAGGATGAGCAAAATCTCATGCA 960  
901 ATCAAGGATGCTGCTTCAAACTTGCATGAGAGGATGAGCAAAATCTCATGCA 960  
961 GCTATTACAGCATGAGAGAGAGGATTTATGACAGTAAGAGAGATGATGATAGAA 1020  
961 GCTATTACAGCATGAGAGAGAGGATTTATGACAGTAAGAGAGATGATGATAGAA 1020  
1021 AGGCTGAATTTCTGTAATTAAGCAACAGCCTGCTTACAGAGAGGATCAATAGAT 1080  
1021 AGGCTGAATTTCTGTAATTAAGCAACAGCCTGCTTACAGAGAGGATCAATAGAT 1080  
1081 GGGCTGGAAGTAAAGCAATGTAATGATGAGGAGTCCAGACAGAAAAAGGATG 1140  
1081 GGGCTGGAAGTAAAGCAATGTAATGATGAGGAGTCCAGACAGAAAAAGGATG 1140  
1141 ATCTGAATGCTGATCCCTGCTGAGAGAGAGAGATGAGATTAAGAGAACTGCAATGCT 1200  
1141 ATCTGAATGCTGATCCCTGCTGAGAGAGAGAGATGAGATTAAGAGAACTGCAATGCT 1200

1201 CAGAGAACTCTAGAGATACAGATGCTCTGAGATTAACCTAAATAGCAGATTCAGA 1260  
1201 CAGAGAACTCTAGAGATACAGATGCTCTGAGATTAACCTAAATAGCAGATTCAGA 1260  
1261 AAGTTAATGAGTGTCTTCCAGAAAGTGAATGAACTTTAGTTGCTGATACACATGATG 1320  
1261 AAGTTAATGAGTGTCTTCCAGAAAGTGAATGAACTTTAGTTGCTGATACACATGATG 1320  
1321 GGGAGTCTGATTCAAATCCCAAGAGTGAATGTTTGGAGCTTCTAAATGAGTATG 1380  
1321 GGGAGTCTGATTCAAATCCCAAGAGTGAATGTTTGGAGCTTCTAAATGAGTATG 1380  
1381 AATATTTCTGTTCTTACAGAGAAATAGACTTACGAGAGATTCCTCAGAGCTTTAA 1440  
1381 AATATTTCTGTTCTTACAGAGAAATAGACTTACGAGAGATTCCTCAGAGCTTTAA 1440  
1441 TATGTAAGAGTGAAGATTCACCTTCAAAATCAGTGAAGATTAATTTGAAGCAAAATAT 1500  
1441 TATGTAAGAGTGAAGATTCACCTTCAAAATCAGTGAAGATTAATTTGAAGCAAAATAT 1500  
1501 TTGGGAAACCTATACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560  
1501 TTGGGAAACCTATACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560  
1561 TAAATATGAGAGATTTGTTACTGAGCAGATTAATCAAGAGAGAGAGAGAGAGAG 1620  
1561 TAAATATGAGAGATTTGTTACTGAGCAGATTAATCAAGAGAGAGAGAGAGAGAG 1620  
1621 AATTAAGCCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680  
1621 AATTAAGCCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680  
1681 CAGATTTGGAGTCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740  
1681 CAGATTTGGAGTCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740  
1741 AGAATGCTCAAGATGATTAATTTACTATGATGCTGATGAGATTAATTAAGAGAG 1800  
1741 AGAATGCTCAAGATGATTAATTTACTATGATGCTGATGAGATTAATTAAGAGAG 1800  
1801 CTATTGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860  
1801 CTATTGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860  
1861 AAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920  
1861 AAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920  
1921 ACAATTTCAAAAGCCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980  
1921 ACAATTTCAAAAGCCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980  
1981 ATGGCTTTGAAGTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2040  
1981 ATGGCTTTGAAGTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2040  
2041 TTGATAGTTTCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2100  
2041 TTGATAGTTTCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2100  
2101 GGCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2160  
2101 GGCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2160  
2161 GTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220  
2161 GTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220  
2221 AGTTAACAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2280  
2221 AGTTAACAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2280  
2281 TTGTCATCTGAGCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2340  
2281 TTGTCATCTGAGCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2340

[illegible]



```

QY 4501 GTGAATACCCATTAAGCCAGATTCAGAGGCCCTTTGTGTCAGAACTTTGAGGTGTCTG 4560
D 4501 GTGAATACCCATTAAGCCAGATTCAGAGGCCCTTTGTGTCAGAACTTTGAGGTGTCTG 4560
QY 4561 CAGATAGCTTACCAGTAAATAAAGAACCCAGAGTGAAGAGTCAATCCCTTTCTAAT 4620
D 4561 CAGATAGCTTACCAGTAAATAAAGAACCCAGAGTGAAGAGTCAATCCCTTTCTAAT 4620
QY 4621 GCCATCATTAAGTAAAGTGGTACATGCACAGTTGCTCTGGAGTCTTCCAGATGA 4680
D 4621 GCCATCATTAAGTAAAGTGGTACATGCACAGTTGCTCTGGAGTCTTCCAGATGA 4680
QY 4681 ACTAACCATTCAAGAGAGCTCATTAAGTTGTTGATGTGAGAGCAACAGCTGGAG 4740
D 4681 ACTAACCATTCAAGAGAGCTCATTAAGTTGTTGATGTGAGAGCAACAGCTGGAG 4740
QY 4741 AGTCTGGGCCACAGATTTGACGGAACATCTTACTTCCCAAGGCAAGATCTAGAGGAA 4800
D 4741 AGTCTGGGCCACAGATTTGACGGAACATCTTACTTCCCAAGGCAAGATCTAGAGGAA 4800
QY 4801 CCCCTTACCTGGAATCTGGAATCAAGCTCTTCTGTGATGACCTGGAATCTGATCCTG 4860
D 4801 CCCCTTACCTGGAATCTGGAATCAAGCTCTTCTGTGATGACCTGGAATCTGATCCTG 4860
QY 4861 AAGACAGAGCCCGAGAGTCAAGCTGCTGTGCAACATACATCTTCAACCTCTGCAATTGA 4920
D 4861 AAGACAGAGCCCGAGAGTCAAGCTGCTGTGCAACATACATCTTCAACCTCTGCAATTGA 4920
QY 4921 AAGTCCCAATTAAGAACTGTCAGAACTCTCCAGAGTCCAGCTGCTGCTCATACTG 4980
D 4921 AAGTCCCAATTAAGAACTGTCAGAACTCTCCAGAGTCCAGCTGCTGCTCATACTG 4980
QY 4981 ATACTGCTGGTAAATGAATGGAAGAGTGTGACAGAGGAGAACCCGAATTTGACAG 5040
D 4981 ATACTGCTGGTAAATGAATGGAAGAGTGTGACAGAGGAGAACCCGAATTTGACAG 5040
QY 5041 CTTCAACAGAAAGGCTCAACAAAAGAAATGTCATGAGTGTGCTGAGCTCAACCCAGAG 5100
D 5041 CTTCAACAGAAAGGCTCAACAAAAGAAATGTCATGAGTGTGCTGAGCTCAACCCAGAG 5100
QY 5101 AATTATGCTGCTGTACAGATTTGCCAGAAAACACCAATCACTTAACTAATTA 5160
D 5101 AATTATGCTGCTGTACAGATTTGCCAGAAAACACCAATCACTTAACTAATTA 5160
QY 5161 CTGAAGAGTACTCAGTGTGTATGAACAGATGCTGAGTTGTGTGGAAGGAGCAC 5220
D 5161 CTGAAGAGTACTCAGTGTGTATGAACAGATGCTGAGTTGTGTGGAAGGAGCAC 5220
QY 5221 TGAATATTTTCTAGCAATTCGCGGAGAAATGGTAGTTAGCTATTTCTGGGTGACC 5280
D 5221 TGAATATTTTCTAGCAATTCGCGGAGAAATGGTAGTTAGCTATTTCTGGGTGACC 5280
QY 5281 AGTCTATTAAGAAAGAAATGCTGTAAGTACATCTTGAAGTCAAGAGAGATGTGG 5340
D 5281 AGTCTATTAAGAAAGAAATGCTGTAAGTACATCTTGAAGTCAAGAGAGATGTGG 5340
QY 5341 TCAATGGAAGAAACCCCAAGGTCGAAGGAGCAAGATCCCAAGAGAAAGATCT 5400
D 5341 TCAATGGAAGAAACCCCAAGGTCGAAGGAGCAAGATCCCAAGAGAAAGATCT 5400
QY 5401 TCAGAGGGCTAGAAATCTGTTGATGGCCCTTACCAACATGCCCAGAGATCAATGG 5460
D 5401 TCAGAGGGCTAGAAATCTGTTGATGGCCCTTACCAACATGCCCAGAGATCAATGG 5460
QY 5461 AATGATGTAACAGCTGTGTGTCTGTGTGCTGAAGAGCTTTCATCAATCAACCTTG 5520
D 5461 AATGATGTAACAGCTGTGTGTCTGTGTGCTGAAGAGCTTTCATCAATCAACCTTG 5520
QY 5521 GCACAGGTGTCAACCCATGTTGTTGTGACGCAAGATGCTGACAGAGAGACATGGCT 5580
D 5521 GCACAGGTGTCAACCCATGTTGTTGTGACGCAAGATGCTGACAGAGAGACATGGCT 5580

```

```

QY 5581 TCCATGCAATTTGGCAGATGTGTGAGGACCCTGTGTGAGCCGAGAGTGGGTGTGGACA 5640
D 5581 TCCATGCAATTTGGCAGATGTGTGAGGACCCTGTGTGAGCCGAGAGTGGGTGTGGACA 5640
QY 5641 GTGTAGCACTTACCAGTGTGCGAGAGCTGGACACTTACTATATCCCAAGATCCCCACA 5700
D 5641 GTGTAGCACTTACCAGTGTGCGAGAGCTGGACACTTACTATATCCCAAGATCCCCACA 5700
QY 5701 GCCACTACTGA 5711
D 5701 GCCACTACTGA 5711

RESULT 8
AAV46459
ID AAV46459 standard; cDNA: 5711 BP.
XX
AC AAV46459;
XX
DT 18-NOV-1998 (first entry)
XX
DE Human BRCA1 omi2 polymorphism #2 cDNA.
XX
KW BRCA1; omi2; human; breast and ovarian cancer predisposing gene;
KW polymorphism; susceptibility; anti-oncogene; tumour suppressor;
KW chromosome 17q; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 120..5711
FT /tag= a
FT /product= "BRCA1 omi2 protein"
FT variation 2430
FT /tag= b
FT /note= "This polymorphic variation can be a T or C
nucleotide"
XX
PN US5750400-A.
XX
PD 12-MAY-1998.
XX
PE 12-FEB-1997; 97US-0798691.
XX
PR 12-FEB-1996; 96US-0598591.
PR 12-FEB-1997; 97US-0798691.
XX
PA (ONCO-) ONCOMED INC.
XX
PI Allen AC, Alvares CP, Critz BS, Murphy PD, Olson SJ;
PI Scheider DB, Zeng B;
XX
DR WPI: 1998-296774/26.
XX
PT BRCA1 omi gene coding sequences - useful for distinguishing between
PT polymorphisms and mutation(s) in the screening for disposition to
PT breast or ovarian cancer
XX
PS Claim 2e; Page -: 54pp; English.
XX
CC This sequence encodes a human BRCA1 (breast and ovarian cancer
CC predisposing gene) omi2 gene in which a polymorphic variation occurs at
CC nucleotide 2430. This sequence and other polymorphic variations of this
CC sequence are useful for the identification of an individual who may or
CC may not have an increased susceptibility to breast or ovarian cancer.
CC The sequences used identify gene changes which are due to polymorphisms,
CC rather than cancer causing mutations. BRCA1 is an anti-oncogene (tumour
CC suppressor) which is involved in genetic inheritance of cancers,
CC especially breast and ovarian cancer. It is found at human chromosome
CC 17q which is known to be linked to cancer susceptibility, especially
CC breast cancer. Cells containing a mutation in this gene lose the
CC wild-type function of BRCA1 and are more susceptible to cancers.
CC
CC NOTE: This sequence does not appear in the specification but has been

```



CC created from the wild type BRCA1 omi2 gene represented in AAV46449.  
XX Sequence 5711 BP; 1956 A; 1098 C; 1274 G; 1382 T; 1 other;  
Query Match 100.0%; Score 5709; DB 19; Length 5711;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 5709; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGCTCGGTGAGACTCTCTGGACCCCGCAGCAGCTGTGGGGTTCTCAGATAACTGGGCC 60  
DB 1 AGCTCGGTGAGACTCTCTGGACCCCGCAGCAGCTGTGGGGTTCTCAGATAACTGGGCC 60  
QY 61 CCGCGCTGAGAGGCTTACCCCTGCTCTGGGTAAGTTCATTTGGAACAGAGAAA 120  
DB 61 CCGCGCTGAGAGGCTTACCCCTGCTCTGGGTAAGTTCATTTGGAACAGAGAAA 120  
QY 121 TGGATTATGCTGCTCGGGTTGAAGAAGTACAAATGTCATTAATGCTATGCAAAA 180  
DB 121 TGGATTATGCTGCTCGGGTTGAAGAAGTACAAATGTCATTAATGCTATGCAAAA 180  
QY 181 TCTTAGAGTGTCCATCTGCTGGAGTTGATCAAGAACTGTCTCCACAAGGTGACC 240  
DB 181 TCTTAGAGTGTCCATCTGCTGGAGTTGATCAAGAACTGTCTCCACAAGGTGACC 240  
QY 241 ACATATTTTGGCAATTTTGCATGCTGAACCTTCAACACAGAGAAAAGGCTTACAGT 300  
DB 241 ACATATTTTGGCAATTTTGCATGCTGAACCTTCAACACAGAGAAAAGGCTTACAGT 300  
QY 301 GTCCCTATGTAAGTATGATATTAACCAAAAGAGCCTACAGAAAGTACAGATTTAGTC 360  
DB 301 GTCCCTATGTAAGTATGATATTAACCAAAAGAGCCTACAGAAAGTACAGATTTAGTC 360  
QY 361 AACTGTTGAGAGCTATTGAAATCATTTGTGCTTTACGCTTGACAGAGTTGGAGT 420  
DB 361 AACTGTTGAGAGCTATTGAAATCATTTGTGCTTTACGCTTGACAGAGTTGGAGT 420  
QY 421 ATGCAAAAGCTATTAATTTTGGCAAAAAGAAATAAATCTCTCTGAAATCTAAAGATG 480  
DB 421 ATGCAAAAGCTATTAATTTTGGCAAAAAGAAATAAATCTCTCTGAAATCTAAAGATG 480  
QY 481 AAGTTCTATCATCCAAAGATGAGGCTACAGAAACGTCGCAAAAGACTTCTACAGAGT 540  
DB 481 AAGTTCTATCATCCAAAGATGAGGCTACAGAAACGTCGCAAAAGACTTCTACAGAGT 540  
QY 541 AACCCGAAATCCCTCTGTCAGAGAAACAGTCTCACTGTCAACCTCTTAACCTTGAA 600  
DB 541 AACCCGAAATCCCTCTGTCAGAGAAACAGTCTCACTGTCAACCTCTTAACCTTGAA 600  
QY 601 CTGTGGAAGTCTGAGAGCAAAAGAGGATACAACTCAAAAGACCTGTCTACATTTG 660  
DB 601 CTGTGGAAGTCTGAGAGCAAAAGAGGATACAACTCAAAAGACCTGTCTACATTTG 660  
QY 661 AATGAGATCTGATCTTCTGGAAGTACCGTTAATAGGCACTTATGAGTGGAG 720  
DB 661 AATGAGATCTGATCTTCTGGAAGTACCGTTAATAGGCACTTATGAGTGGAG 720  
QY 721 ATCAGAAATTTGTAACAATCACCCCTCAAGAGAACAGGATGAATCAGTTGGATCTG 780  
DB 721 ATCAGAAATTTGTAACAATCACCCCTCAAGAGAACAGGATGAATCAGTTGGATCTG 780  
QY 781 CAAAAAAGGCTGCTTGAATTTTCTGAGACGATGTAACTAATGATGAGATCATCAAC 840  
DB 781 CAAAAAAGGCTGCTTGAATTTTCTGAGACGATGTAACTAATGATGAGATCATCAAC 840  
QY 841 CGAGTAATATGATTTGAACACACGTGAGAGGTCAGCTGAGAGCATCCAGAAAAGT 900  
DB 841 CGAGTAATATGATTTGAACACACGTGAGAGGTCAGCTGAGAGCATCCAGAAAAGT 900  
QY 901 ATCAGAGTATGTTCTGTTTCAAACTTGATGAGGATGTGGACAAATATCATGCA 960  
DB 901 ATCAGAGTATGTTCTGTTTCAAACTTGATGAGGATGTGGACAAATATCATGCA 960  
QY 961 GCTCATATTCAGCATGAGAACAGCATTTTATTAATCACTAAAGACAGATGAATGTAGAAA 1020

DB 961 GCTCATATTCAGCATGAGAACAGCATTTTATTAATCACTAAAGACAGAAATGATAGAAA 1020  
QY 1021 AGCTGAAATCTGTAATAAAGCAANAGGCTGGCTTAGCAGAGACCAATACAGAT 1080  
DB 1021 AGCTGAAATCTGTAATAAAGCAANAGGCTGGCTTAGCAGAGACCAATACAGAT 1080  
QY 1081 GGGCTGGAAGTAAAGCAACATGTAATGATAGGCGGACCTCCACACAGAAAAAGGTG 1140  
DB 1081 GGGCTGGAAGTAAAGCAACATGTAATGATAGGCGGACCTCCACACAGAAAAAGGTG 1140  
QY 1141 ATCTGAATGCTGATCCCTGTGTGAGAGAAAGATGGAATTAAGCAGAACTGCCATGCT 1200  
DB 1141 ATCTGAATGCTGATCCCTGTGTGAGAGAAAGATGGAATTAAGCAGAACTGCCATGCT 1200  
QY 1201 CAGGAATCCTAGAGATGACTGGAAGATGCTGCTGGATTAACACTAAATATGACATTCACA 1260  
DB 1201 CAGGAATCCTAGAGATGACTGGAAGATGCTGCTGGATTAACACTAAATATGACATTCACA 1260  
QY 1261 AAGTTAATGAGTGTGTTCCAGAAAGTATGAACCTGTTAGGTTCTGATGACTCAGATGATG 1320  
DB 1261 AAGTTAATGAGTGTGTTCCAGAAAGTATGAACCTGTTAGGTTCTGATGACTCAGATGATG 1320  
QY 1321 GGGAGTCTGAATCAATGCCAAAGTACCTGATTAATGACGTTCAATGAGTATGATG 1380  
DB 1321 GGGAGTCTGAATCAATGCCAAAGTACCTGATTAATGACGTTCAATGAGTATGATG 1380  
QY 1381 AATATCTGCTCTTCAGAAAAATTAAGTACTGAGCCAGTGTCTCATGAGGCTTTAA 1440  
DB 1381 AATATCTGCTCTTCAGAAAAATTAAGTACTGAGCCAGTGTCTCATGAGGCTTTAA 1440  
QY 1441 TATGTAAGTGAAGAGTTCACCTCAAAATCACTAAGATATATTTGAAGCAAAATAT 1500  
DB 1441 TATGTAAGTGAAGAGTTCACCTCAAAATCACTAAGATATATTTGAAGCAAAATAT 1500  
QY 1501 TTGGGAAACCTATTCGAGAGAGGCAAGCCTCCCACTTAACCCATGTAAGTGAATATC 1560  
DB 1501 TTGGGAAACCTATTCGAGAGAGGCAAGCCTCCCACTTAACCCATGTAAGTGAATATC 1560  
QY 1561 TAAATATGAGGATTTGTTACTGAGCCAGATTAATGATTAAGAGGCTCCCTCACAAATA 1620  
DB 1561 TAAATATGAGGATTTGTTACTGAGCCAGATTAATGATTAAGAGGCTCCCTCACAAATA 1620  
QY 1621 AATTAAGCGTAAAGAGACCTACATCAGGCTTCATCCTGAGATTTTATCAAGAAAG 1680  
DB 1621 AATTAAGCGTAAAGAGACCTACATCAGGCTTCATCCTGAGATTTTATCAAGAAAG 1680  
QY 1681 CAGATTTGGCAGTTCAAAAGACTCTCGAATATGTAATCAGGGAACCTAACCAAGCGAGC 1740  
DB 1681 CAGATTTGGCAGTTCAAAAGACTCTCGAATATGTAATCAGGGAACCTAACCAAGCGAGC 1740  
QY 1741 AGAATGGTCAAGTATGATTAATTAATAGTGGTATGATGATGAATTAACCAAAAGGTGAT 1800  
DB 1741 AGAATGGTCAAGTATGATTAATTAATAGTGGTATGATGATGAATTAACCAAAAGGTGAT 1800  
QY 1801 CTATTCAGATGAGAAAAATCTTAACCCATAGAAATCACTCCAAAAAGATCTGTTTCA 1860  
DB 1801 CTATTCAGATGAGAAAAATCTTAACCCATAGAAATCACTCCAAAAAGATCTGTTTCA 1860  
QY 1861 AAAAGAAAGCTGAACCTATTAAGCAGATTAATGATGATGATGATGATGATGATGATG 1920  
DB 1861 AAAAGAAAGCTGAACCTATTAAGCAGATTAATGATGATGATGATGATGATGATGATG 1920  
QY 1921 ACAATTCAAAAGCACTTAAAGAAATAGGCTGAGAGGGAAGTCTTCTACAGAGCATATTC 1980  
DB 1921 ACAATTCAAAAGCACTTAAAGAAATAGGCTGAGAGGGAAGTCTTCTACAGAGCATATTC 1980  
QY 1981 ATGCGCTTGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2040  
DB 1981 ATGCGCTTGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2040  
QY 2041 TTGATAGTGTGTTCTAGCAGTGAAGATTAAGAAAAAAGTACAAACCAATGCGAGTCA 2100  
DB 2041 TTGATAGTGTGTTCTAGCAGTGAAGATTAAGAAAAAAGTACAAACCAATGCGAGTCA 2100

Db	2041	TTGATAGTTGTTTACGAGTGAAGAGATTAAGAAAAAAGTACACACCAATTCGCACTCA	2100
OY	2101	GGCAGCAGCAAACTTACACTCATGTGAAAGGTAAAGAACTGCAACTGAGAGCCAAAGA	2160
Db	2101	GGCAGCAGCAAACTTACACTCATGTGAAAGGTAAAGAACTGCAACTGAGAGCCAAAGA	2160
OY	2161	GTATCAAGCCAAATGAAACAGACAGTAAAGACATGACAGTGAATCTTCCAGAGCTGA	2220
Db	2161	GTATCAAGCCAAATGAAACAGACAGTAAAGACATGACAGTGAATCTTCCAGAGCTGA	2220
OY	2221	AGTTAAACAATATGACCTGGTCTTTTCTAAAGTTCCTAAATACAGAGAACTTAAAGAT	2280
Db	2221	AGTTAAACAATATGACCTGGTCTTTTCTAAAGTTCCTAAATACAGAGAACTTAAAGAT	2280
OY	2281	TTGTCAATCCTAGCCCTTCCAAGAGAAAGAAAGAACTAGAAACAGATTAAAGTG	2340
Db	2281	TTGTCAATCCTAGCCCTTCCAAGAGAAAGAAAGAACTAGAAACAGATTAAAGTG	2340
OY	2341	CTAATATATGCTGAAGACCCCAAAGATCTCATGTTAAGTGAAGAAAGGTTTGCAACTG	2400
Db	2341	CTAATATATGCTGAAGACCCCAAAGATCTCATGTTAAGTGAAGAAAGGTTTGCAACTG	2400
OY	2401	AAAGATCTGTAGAGATGACAGTATTTTCATGTGGTACCGAGTACGATTAATGAGCCTAG	2460
Db	2401	AAAGATCTGTAGAGATGACAGTATTTTCATGTGGTACCGAGTACGATTAATGAGCCTAG	2460
OY	2461	AAAGATCTCTTACTGTGAAGTATGACACTCTAGGGAAGAGCCAAAACAGAACCAATAAAT	2520
Db	2461	AAAGATCTCTTACTGTGAAGTATGACACTCTAGGGAAGAGCCAAAACAGAACCAATAAAT	2520
OY	2521	GTGTGAGTCACTGTGACAGCATTTGAAAAACCCCAAGGAGCTAATTCATGGTTGTTCCAAG	2580
Db	2521	GTGTGAGTCACTGTGACAGCATTTGAAAAACCCCAAGGAGCTAATTCATGGTTGTTCCAAG	2580
OY	2581	ATATATAGAAATGACACAGAAAGGCTTTAAGTTCATTTGGGACAGAGATTAACACAGTC	2640
Db	2581	ATATATAGAAATGACACAGAAAGGCTTTAAGTTCATTTGGGACAGAGATTAACACAGTC	2640
OY	2641	GGGAACAGAGCATAGAAATGGAAGAAAGTAGTAATGATGCTCAGTATTTGGCAGAAATCAT	2700
Db	2641	GGGAACAGAGCATAGAAATGGAAGAAAGTAGTAATGATGCTCAGTATTTGGCAGAAATCAT	2700
OY	2701	TCAAGGTTTCAAGAGCCGACATTTGCTGCTGTTTCAAAATCCAGAAATGACAGAAAGG	2760
Db	2701	TCAAGGTTTCAAGAGCCGACATTTGCTGCTGTTTCAAAATCCAGAAATGACAGAAAGG	2760
OY	2761	AATGTGCAACATCTCTGCCCCACCTCTGGGTCTTAAAGAAACAAAGTCCAAAGTCACTT	2820
Db	2761	AATGTGCAACATCTCTGCCCCACCTCTGGGTCTTAAAGAAACAAAGTCCAAAGTCACTT	2820
OY	2821	TTGATATGGAACAAGAAAGAAATATGAGAAAGATGAGTCTTAATATCAAGCTGTAC	2880
Db	2821	TTGATATGGAACAAGAAAGAAATATGAGAAAGATGAGTCTTAATATCAAGCTGTAC	2880
OY	2881	AGACAGTTAATATACATGACAGAGCTTTCCTGTGTGGTCAAGAAAGATTAAGCCAGTTGATA	2940
Db	2881	AGACAGTTAATATACATGACAGAGCTTTCCTGTGTGGTCAAGAAAGATTAAGCCAGTTGATA	2940
OY	2941	ATGCCAAATATGATATCAAGAGAGGCTCTAGGTTTGTCTATCTATCTCAGTTTCAGAGCA	3000
Db	2941	ATGCCAAATATGATATCAAGAGAGGCTCTAGGTTTGTCTATCTATCTCAGTTTCAGAGCA	3000
OY	3001	ACGAAACCTGGAGCTATATACCAATTAACATGAGACTTTTACAAACCCCATATCGTATAC	3060
Db	3001	ACGAAACCTGGAGCTATATACCAATTAACATGAGACTTTTACAAACCCCATATCGTATAC	3060
OY	3061	CACCACTTTTCCCATCAAGTCAATTTGTTAAACATAATGTAAAGAAAAATCTGTTAGAG	3120
Db	3061	CACCACTTTTCCCATCAAGTCAATTTGTTAAACATAATGTAAAGAAAAATCTGTTAGAG	3120
OY	3121	AAAACTTTGAGGAACATTCAATGTCTACCTGAAAGAGAAATGGAATGAGACATTTCCAA	3180
Db	3121	AAAACTTTGAGGAACATTCAATGTCTACCTGAAAGAGAAATGGAATGAGACATTTCCAA	3180

OY	3181	GTACAGTGTAGACACATTTAGCCGTATAATCACTTAAGAGAAATGTTTTAAGAACCCAGCT	3240
Db	3181	GTACAGTGTAGACACATTTAGCCGTATAATCACTTAAGAGAAATGTTTTAAGAACCCAGCT	3240
OY	3241	CACCAATATTATATGAAGTAGTGTCCACTACTAATGAAGTGGGCCCGCTATTATATGANA	3300
Db	3241	CACCAATATTATATGAAGTAGTGTCCACTACTAATGAAGTGGGCCCGCTATTATATGANA	3300
OY	3301	TAGCTTTCAGTATGAAAACATTCAGCAGAACTAGTAGAAGAACGAGGGCCAAATTGA	3360
Db	3301	TAGCTTTCAGTATGAAAACATTCAGCAGAACTAGTAGAAGAACGAGGGCCAAATTGA	3360
OY	3361	ATGCTATGCTTAGATTAGGGGTTTTGCCAAGCTGATATAAACAAAGCTTCTCTGGAA	3420
Db	3361	ATGCTATGCTTAGATTAGGGGTTTTGCCAAGCTGATATAAACAAAGCTTCTCTGGAA	3420
OY	3421	GTAATTTGAACCATTCCTGAAATATAAAAAAGCAAGTAATGAAGAGTAGTTCAGACTGTA	3480
Db	3421	GTAATTTGAACCATTCCTGAAATATAAAAAAGCAAGTAATGAAGAGTAGTTCAGACTGTA	3480
OY	3481	ATTACAGATTTCTCCATATCTGATTTGCATTACTTAGAAGACACCTATGGAGTAGTACT	3540
Db	3481	ATTACAGATTTCTCCATATCTGATTTGCATTACTTAGAAGACACCTATGGAGTAGTACT	3540
OY	3541	ATGCATCTCAGGTTGTTCTTGAGAGACACCTGATGACCTGTTAGATGATGGTGAATAAAG	3600
Db	3541	ATGCATCTCAGGTTGTTCTTGAGAGACACCTGATGACCTGTTAGATGATGGTGAATAAAG	3600
OY	3601	AAGATTAAGTTTGTGCTGAAAATATGACATTAAGGAAAGTTCTGCTGTTTTAGCAAAACCG	3660
Db	3601	AAGATTAAGTTTGTGCTGAAAATATGACATTAAGGAAAGTTCTGCTGTTTTAGCAAAACCG	3660
OY	3661	TCCAGAAAGAGAGACTTTCAGAGAGTCTTACCCCTTACCCATACACATTTGGCTCAGG	3720
Db	3661	TCCAGAAAGAGAGACTTTCAGAGAGTCTTACCCCTTACCCATACACATTTGGCTCAGG	3720
OY	3721	GTTACCGAAGAGGGGGCCAGAATAATTAGAGTCCCTGAGAGAGAACTTATCTAGTAGATG	3780
Db	3721	GTTACCGAAGAGGGGGCCAGAATAATTAGAGTCCCTGAGAGAGAACTTATCTAGTAGATG	3780
OY	3781	AAGAGCTTCCCTGCTTCCACACCTTGTATTGTTGTTAAAGTAAACATATACCTTCTCAGT	3840
Db	3781	AAGAGCTTCCCTGCTTCCACACCTTGTATTGTTGTTAAAGTAAACATATACCTTCTCAGT	3840
OY	3841	CTACTAGGCATAGACACCGTGTCTTACCGAGTGTCTGTCTTAAGACACAGAGAGAAATTTAT	3900
Db	3841	CTACTAGGCATAGACACCGTGTCTTACCGAGTGTCTGTCTTAAGACACAGAGAGAAATTTAT	3900
OY	3901	TATCATTTGAAGAAATAGCTTAAATGACTGCTAGTAACACCAAGTATATTTGGCAAGGCATCTC	3960
Db	3901	TATCATTTGAAGAAATAGCTTAAATGACTGCTAGTAACACCAAGTATATTTGGCAAGGCATCTC	3960
OY	3961	AGGAACATCACCTTAGTGTAGGAAAACAAATGTCTGCTAGCTTGTCTTCTCACAGTGTCA	4020
Db	3961	AGGAACATCACCTTAGTGTAGGAAAACAAATGTCTGCTAGCTTGTCTTCTCACAGTGTCA	4020
OY	4021	GTTGAATTTGGAAGACTTGAATGCTGCAAAATATAAACACCCAGAGATCTTCTTGAATTTGGTCTT	4080
Db	4021	GTTGAATTTGGAAGACTTGAATGCTGCAAAATATAAACACCCAGAGATCTTCTTGAATTTGGTCTT	4080
OY	4081	CCAACAAATATAGCATCAGTCTGTAAGGCCAGGAGTTGGCTGTAAGTACCAAGGAATTTGG	4140
Db	4081	CCAACAAATATAGCATCAGTCTGTAAGGCCAGGAGTTGGCTGTAAGTACCAAGGAATTTGG	4140
OY	4141	TTTCAGATGATGAAGAAAGAGAGAAAGGCGTTGGAAGAAAATATATCAAGAAAGCAAAAGCA	4200
Db	4141	TTTCAGATGATGAAGAAAGAGAGAAAGGCGTTGGAAGAAAATATATCAAGAAAGCAAAAGCA	4200
OY	4201	TGGATTCAAACTTATAGTGTGAAGCAGACATCTGGGTTGTAGAGTGAAGAAACAGCTCTCTGAG	4260
Db	4201	TGGATTCAAACTTATAGTGTGAAGCAGACATCTGGGTTGTAGAGTGAAGAAACAGCTCTCTGAG	4260

QY 4261 ACTGCTCAGGGCTATCCTCTCAGAGTGACATTTTAACTCCTCAGCAGAGGGATACCATGC 4320  
 |||||  
 Db 4261 ACTGCTCAGGGCTATCCTCTCAGAGTGACATTTTAACTCCTCAGCAGAGGGATACCATGC 4320  
 QY 4321 AACATACCTGATTAAGCTCCAGAGGAATGGCTGACTAGAACTGTGTAGACAGC 4380  
 |||||  
 Db 4321 AACATACCTGATTAAGCTCCAGAGGAATGGCTGACTAGAACTGTGTAGACAGC 4380  
 QY 4381 ATGGAGGCCAGCTTCTAACAGGTACCTCCATCATTAAGTGCATCTCTGCCCCCTTAGG 4440  
 |||||  
 Db 4381 ATGGAGGCCAGCTTCTAACAGGTACCTCCATCATTAAGTGCATCTCTGCCCCCTTAGG 4440  
 QY 4441 ACCTGCGAAATCCAGAAACCAAGACATCAGAAAAAGCAGTATTAACTTCACAGAAAAATA 4500  
 |||||  
 Db 4441 ACCTGCGAAATCCAGAAACCAAGACATCAGAAAAAGCAGTATTAACTTCACAGAAAAATA 4500  
 QY 4501 GTGAATACCTTAATAGCCAAATCCAGAAAGCCCTTCTGCTGACAACTTTGAGGTCTG 4560  
 |||||  
 Db 4501 GTGAATACCTTAATAGCCAAATCCAGAAAGCCCTTCTGCTGACAACTTTGAGGTCTG 4560  
 QY 4561 CAGATAGTTCTACAGTAAATAAAGAACAGAGAGTCAATCCCTCTTAAT 4620  
 |||||  
 Db 4561 CAGATAGTTCTACAGTAAATAAAGAACAGAGAGTCAATCCCTCTTAAT 4620  
 QY 4621 GCCCATCATTAATGATGATAGTGTGACATGCACAGTTGCTGCGAGTCTTCAGAAATAGAA 4680  
 |||||  
 Db 4621 GCCCATCATTAATGATGATAGTGTGACATGCACAGTTGCTGCGAGTCTTCAGAAATAGAA 4680  
 QY 4681 ACTACCATCTCAAGAGAGAGCTCATTAAAGTTGTGATGTGAGAGACCAAGCTGGAAG 4740  
 |||||  
 Db 4681 ACTACCATCTCAAGAGAGAGCTCATTAAAGTTGTGATGTGAGAGACCAAGCTGGAAG 4740  
 QY 4741 AGTCTGGGCCACACGATTTTGACGAAACATCTTACTTGCCAGAGCAAGATCTAGAGGAA 4800  
 |||||  
 Db 4741 AGTCTGGGCCACACGATTTTGACGAAACATCTTACTTGCCAGAGCAAGATCTAGAGGAA 4800  
 QY 4801 CCCCTTACCTGGAATGTGGAATCAGCCTCTCTGTGATGACCTGAACTGATCTTCTG 4860  
 |||||  
 Db 4801 CCCCTTACCTGGAATGTGGAATCAGCCTCTCTGTGATGACCTGAACTGATCTTCTG 4860  
 QY 4861 AAGACAGAGCCCCAGAGTCACTGCTGTGGAACATACCATCTTCAACCTCTGCAATGGA 4920  
 |||||  
 Db 4861 AAGACAGAGCCCCAGAGTCACTGCTGTGGAACATACCATCTTCAACCTCTGCAATGGA 4920  
 QY 4921 AAGTCCCAATTTGAAAGTTGCAAGATCTGCCCCAGAGTCCAGCTGCTCATACTAGT 4980  
 |||||  
 Db 4921 AAGTCCCAATTTGAAAGTTGCAAGATCTGCCCCAGAGTCCAGCTGCTCATACTAGT 4980  
 QY 4981 ATACTGCTGGGTATTAATGCAATGGAAGAAAGTGTGAGCAGAGGAGAACCCAGAAATGACAG 5040  
 |||||  
 Db 4981 ATACTGCTGGGTATTAATGCAATGGAAGAAAGTGTGAGCAGAGGAGAACCCAGAAATGACAG 5040  
 QY 5041 CTTCAACAGAAAGGTCACAAAGAAATGTCATGTGTGCTGAGCTGAGCTGAGCCAGAG 5100  
 |||||  
 Db 5041 CTTCAACAGAAAGGTCACAAAGAAATGTCATGTGTGCTGAGCTGAGCTGAGCCAGAG 5100  
 QY 5101 AATTATGCTCGTGTACAGTTTGCCAGAAACACCAATCACTTAACTAATCTAATTA 5160  
 |||||  
 Db 5101 AATTATGCTCGTGTACAGTTTGCCAGAAACACCAATCACTTAACTAATCTAATTA 5160  
 QY 5161 CTGAAGAGACTACTCATGTTGTATGAAACAGATCTGAGTTGTGTGAAACGGGACAC 5220  
 |||||  
 Db 5161 CTGAAGAGACTACTCATGTTGTATGAAACAGATCTGAGTTGTGTGAAACGGGACAC 5220  
 QY 5221 TGAATAATTTTCTAGAAATTTGCGGAGGAAATGGGAGTATCTATTTTGGGTGACCC 5280  
 |||||  
 Db 5221 TGAATAATTTTCTAGAAATTTGCGGAGGAAATGGGAGTATCTATTTTGGGTGACCC 5280  
 QY 5281 ACTCTAATTAAGAAAGAAATGCTGAATGAGCATATTTTGAAGTCAGAGGAGATGTGG 5340  
 |||||  
 Db 5281 ACTCTAATTAAGAAAGAAATGCTGAATGAGCATATTTTGAAGTCAGAGGAGATGTGG 5340  
 QY 5341 TCAATGGAAGAAACCAACCAAGGTCCAAAGCGAGCAAGAGATCCAGGACGAAAGATCT 5400

Db 5341 TCAATGGAAGAAACCAACCAAGGTCCAAAGCGAGCAAGAGATCCAGAGCAAGAAATCT 5400  
 |||||  
 QY 5401 TCAGGGGGCTAAGAAATCTGTGCTATGAGGCGCCCTTCACCAACATGCCCAAGATCAACTGG 5460  
 |||||  
 Db 5401 TCAGGGGGCTAAGAAATCTGTGCTATGAGGCGCCCTTCACCAACATGCCCAAGATCAACTGG 5460  
 QY 5461 AATGATGATGACAGCTGTGTGCTGCTGCTGTGATGAGAGAGCTTTCATCATTCACCCTTG 5520  
 |||||  
 Db 5461 AATGATGATGACAGCTGTGTGCTGCTGCTGTGATGAGAGAGCTTTCATCATTCACCCTTG 5520  
 QY 5521 GCACAGGTGTCACCAATTTGTTGTGTCAGCCAGATGCTGTGACAGAGCAATGCT 5580  
 |||||  
 Db 5521 GCACAGGTGTCACCAATTTGTTGTGTCAGCCAGATGCTGTGACAGAGCAATGCT 5580  
 QY 5581 TCCATGCAATTTGGGCGAGATGTGTGAGGCACTGTGTGTGACCCGAGAGTGGTGGACA 5640  
 |||||  
 Db 5581 TCCATGCAATTTGGGCGAGATGTGTGAGGCACTGTGTGTGACCCGAGAGTGGTGGACA 5640  
 QY 5641 GTGTAGCACTCTACAGTGCAGAGCTGACACACTACCTGATTAACCCAGATCCCCACA 5700  
 |||||  
 Db 5641 GTGTAGCACTCTACAGTGCAGAGCTGACACACTACCTGATTAACCCAGATCCCCACA 5700  
 QY 5701 GCCACTACTGA 5711  
 |||||  
 Db 5701 GCCACTACTGA 5711  
 |||||  
 RESULT 9  
 AAV46460  
 ID AAV46460 standard; cDNA; 5711 BP.  
 XX  
 AC AAV46460;  
 XX  
 DT 18-NOV-1998 (first entry)  
 XX  
 DE Human BRCA1 omi2 polymorphism #3 cDNA.  
 XX  
 KW BRCA1; omi2; human; breast and ovarian cancer predisposing gene;  
 KW polymorphism; susceptibility; anti-oncogene; tumour suppressor;  
 KW chromosome 17q; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FT Key Location/Qualifiers  
 FT CDS 120..5711  
 FT FT /\*tag= a  
 FT FT /product= "BRCA1 omi2 protein"  
 FT FT variation 2731  
 FT FT /\*tag= b  
 FT FT /note= "This polymorphic variation can be a C or T  
 nucleotide"  
 XX  
 PN US5750400-A.  
 XX  
 PD 12-MAY-1998.  
 XX  
 PF 12-FEB-1997; 97US-0798691.  
 XX  
 PR 12-FEB-1996; 96US-0598591.  
 PR 12-FEB-1997; 97US-0798691.  
 XX  
 PA (ONCO-) ONCORMED INC.  
 XX  
 PI Allen AC, Alvaras CP, Critz BS, Murphy PD, Olson SJ;  
 PI Schelter DB, Zeng B;  
 XX  
 DR WPI; 1998-296774/26.  
 XX  
 PT BRCA1 omi gene coding sequences - useful for distinguishing between  
 PT polymorphisms and mutation(s) in the screening for disposition to  
 PT breast or ovarian cancer  
 XX

PS Claim 2e; Page -: 54pp; English.  
 XX This sequence encodes a human BRCA1 (breast and ovarian cancer  
 CC predisposing gene) om12 gene in which a polymorphic variation occurs at  
 CC nucleotide 2731. This sequence and other polymorphic variations of this  
 CC sequence are useful for the identification of an individual who may or  
 CC may not have an increased susceptibility to breast or ovarian cancer.  
 CC The sequences used identify gene changes which are due to polymorphisms,  
 CC rather than cancer causing mutations. BRCA1 is an anti-oncogene (tumour  
 CC suppressor) which is involved in genetic inheritance of cancers,  
 CC especially breast and ovarian cancer. It is found at human chromosome  
 CC 17q which is known to be linked to cancer susceptibility, especially  
 CC breast cancer. Cells containing a mutation in this gene lose the  
 CC wild-type function of BRCA1 and are more susceptible to cancers.  
 CC NOTE: This sequence does not appear in the specification but has been  
 CC created from the wild type BRCA1 om12 gene represented in AAV6449.  
 XX  
 SQ Sequence 5711 BP; 1956 A; 1098 C; 1274 G; 1382 T; 1 other;  
 Query Match 100.0%; Score 5709; DB 19; Length 5711;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 5709; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 1 ACCTCGCTGAGACTTCTCTGAGACCCGACAGGCTGTGGGTTTTCAGATTAAGTGGCC 60  
 1 AGCTCGCTGAGACTTCTCTGAGACCCGACAGGCTGTGGGTTTTCAGATTAAGTGGCC 60  
 61 CCGGCGCTGAGAGGCTTACCCCTGCTGGGTAAAGTGAATGAGACAGAAAGAAA 120  
 61 CCGGCGCTGAGAGGCTTACCCCTGCTGGGTAAAGTGAATGAGACAGAAAGAAA 120  
 121 TGAATTTATCTGCTCTGCGGTTGAAGAAATGATTAATGATGATGAGAAA 180  
 121 TGAATTTATCTGCTCTGCGGTTGAAGAAATGATTAATGATGATGAGAAA 180  
 121 TGAATTTATCTGCTCTGCGGTTGAAGAAATGATTAATGATGATGAGAAA 180  
 121 TGAATTTATCTGCTCTGCGGTTGAAGAAATGATTAATGATGATGAGAAA 180  
 181 TCTTAGAGTCTCCATCTGTCTGAGTGTATCAAGAAAGTGTCTCCAAAGAGTGTACC 240  
 181 TCTTAGAGTCTCCATCTGTCTGAGTGTATCAAGAAAGTGTCTCCAAAGAGTGTACC 240  
 241 ACATATTTTGAATTTTGTGATGCTGAACTGCAACCAAGAAAGGAGGCTTTCACAGT 300  
 241 ACATATTTTGAATTTTGTGATGCTGAACTGCAACCAAGAAAGGAGGCTTTCACAGT 300  
 241 ACATATTTTGAATTTTGTGATGCTGAACTGCAACCAAGAAAGGAGGCTTTCACAGT 300  
 301 GTCTTTATGTAAGATGATATACCAAAAGAGGCTTACAAAGAAAGTGTATGATGTC 360  
 301 GTCTTTATGTAAGATGATATACCAAAAGAGGCTTACAAAGAAAGTGTATGATGTC 360  
 301 GTCTTTATGTAAGATGATATACCAAAAGAGGCTTACAAAGAAAGTGTATGATGTC 360  
 361 AACCTTTGTAAGAGTATGAAATCATTTTGTCTTTTACGTTGACACAGGTTTGGAGT 420  
 361 AACCTTTGTAAGAGTATGAAATCATTTTGTCTTTTACGTTGACACAGGTTTGGAGT 420  
 421 ATGCAAAACAGCTTAATTTTGTGCAAAAAAGAAATTAATCTCTGAAACATCTAAAGATG 480  
 421 ATGCAAAACAGCTTAATTTTGTGCAAAAAAGAAATTAATCTCTGAAACATCTAAAGATG 480  
 421 ATGCAAAACAGCTTAATTTTGTGCAAAAAAGAAATTAATCTCTGAAACATCTAAAGATG 480  
 481 AAGTTTCTATCATCAAAAGTATGGGTACAGAAACCGTGCAGAAAGCTTCTACAGAGTG 540  
 481 AAGTTTCTATCATCAAAAGTATGGGTACAGAAACCGTGCAGAAAGCTTCTACAGAGTG 540  
 481 AAGTTTCTATCATCAAAAGTATGGGTACAGAAACCGTGCAGAAAGCTTCTACAGAGTG 540  
 541 AACCCGAAATCTTCTCTGAGAAACAGCTCAAGTCTCAACTCTCAACCTTGGAA 600  
 541 AACCCGAAATCTTCTCTGAGAAACAGCTCAAGTCTCAACTCTCAACCTTGGAA 600  
 601 CTGTGAGACTGTGAGAGCAAAAGCAGGATACAACTCAAAAGAGCTGTCTACATG 660  
 601 CTGTGAGACTGTGAGAGCAAAAGCAGGATACAACTCAAAAGAGCTGTCTACATG 660  
 661 AATTTGAGATCTGATTTCTTGAAGATACGTTAATAAGGCAACTTATGCAAGTGTGGAG 720  
 661 AATTTGAGATCTGATTTCTTGAAGATACGTTAATAAGGCAACTTATGCAAGTGTGGAG 720  
 721 ATCAAGATTTGTACAAATACCCCTCAGAGAAACAGGATGAATCAGTTGGATTTCTG 780  
 721 ATCAAGATTTGTACAAATACCCCTCAGAGAAACAGGATGAATCAGTTGGATTTCTG 780

DB 721 ATCAAGATTTGTACAAATACCCCTCAGAGAAACAGGATGAATCAGTTGGATTTCTG 780  
 QY 781 CAAAAAAGGCTGCTGTGAATTTTCTGAGACGATGTAAACAAATACTGAACATCATCAGC 840  
 DB 781 CAAAAAAGGCTGCTGTGAATTTTCTGAGACGATGTAAACAAATACTGAACATCATCAGC 840  
 QY 841 CCAATTAATATGATTTGAGACACCACTGAGAAAGCTGTGAGAGGATTCAGAAAAAGT 900  
 DB 841 CCAATTAATATGATTTGAGACACCACTGAGAAAGCTGTGAGAGGATTCAGAAAAAGT 900  
 QY 901 ATCAGGATGCTGCTTTTCAAACTGATGTGAGAGGATGTGAGACAAATCTCATGCA 960  
 DB 901 ATCAGGATGCTGCTTTTCAAACTGATGTGAGAGGATGTGAGACAAATCTCATGCA 960  
 QY 961 GCTATTACAGCATGAGACAGCGTTTATTTACTACATAAAGACAGATGAATGTAGAAA 1020  
 DB 961 GCTATTACAGCATGAGACAGCGTTTATTTACTACATAAAGACAGATGAATGTAGAAA 1020  
 QY 1021 AGGCTGAATTCCTGTAATTAAGCAACAGGCTGTGAGAGGAGGCAACATTAACAGAT 1080  
 DB 1021 AGGCTGAATTCCTGTAATTAAGCAACAGGCTGTGAGAGGAGGCAACATTAACAGAT 1080  
 QY 1081 GGGCTGGAAGTAAAGAAACATGTATAGTATGAGGAGTCCACAGCAAAAAAGGTAG 1140  
 DB 1081 GGGCTGGAAGTAAAGAAACATGTATAGTATGAGGAGTCCACAGCAAAAAAGGTAG 1140  
 QY 1141 ATCTGAATGCTGATCCCTGTGTGAGAAAGAAAGATGAATGAACCAAACTGCTAGCT 1200  
 DB 1141 ATCTGAATGCTGATCCCTGTGTGAGAAAGAAAGATGAATGAACCAAACTGCTAGCT 1200  
 QY 1201 CAGAGAAATCTAGAGATCTGAGATGATGCTTGTGATTAACATTAATAGCAGCTTACA 1260  
 DB 1201 CAGAGAAATCTAGAGATCTGAGATGATGCTTGTGATTAACATTAATAGCAGCTTACA 1260  
 QY 1261 AAGTTAATGAGTGTGTTTCCAGAAAGTATGATGATGATGATGATGATGATGATG 1320  
 DB 1261 AAGTTAATGAGTGTGTTTCCAGAAAGTATGATGATGATGATGATGATGATGATG 1320  
 QY 1321 GGGAGTGTGAATCAAAAGTGAAGGATGATGATGATGATGATGATGATGATGATG 1380  
 DB 1321 GGGAGTGTGAATCAAAAGTGAAGGATGATGATGATGATGATGATGATGATGATG 1380  
 QY 1381 AATATTTCTGTTTCTGAGAGAAATAGACTTACTGAGGCAATCTCATGAGGCTTTAA 1440  
 DB 1381 AATATTTCTGTTTCTGAGAGAAATAGACTTACTGAGGCAATCTCATGAGGCTTTAA 1440  
 QY 1441 TATGTAAAGTGAAGAGTTCACCTCAATCAGTAGAGATTAATTTGAAGACAAATAT 1500  
 DB 1441 TATGTAAAGTGAAGAGTTCACCTCAATCAGTAGAGATTAATTTGAAGACAAATAT 1500  
 QY 1501 TTTGGAAACCTTATCGGAAAGGCAAGGCTCCCAACTTAAGGATGTAATCTGAATC 1560  
 DB 1501 TTTGGAAACCTTATCGGAAAGGCAAGGCTCCCAACTTAAGGATGTAATCTGAATC 1560  
 QY 1561 TAAATTTATGAGAGATTTTCTGAGAGCAGATTAATTAAGAGGCGCTCCCTACAATA 1620  
 DB 1561 TAAATTTATGAGAGATTTTCTGAGAGCAGATTAATTAAGAGGCGCTCCCTACAATA 1620  
 QY 1621 AATTTAAGGCTTAAAGAGAGCTTACATCAGGCTTCTCATCTGAGAGATTTATCAAG 1680  
 DB 1621 AATTTAAGGCTTAAAGAGAGCTTACATCAGGCTTCTCATCTGAGAGATTTATCAAG 1680  
 QY 1681 CAGATTTGGAGTTCAGAAAGCTCTGAATGTAATTCAGGGAACATCAACCAAGGAGC 1740  
 DB 1681 CAGATTTGGAGTTCAGAAAGCTCTGAATGTAATTCAGGGAACATCAACCAAGGAGC 1740  
 QY 1741 AGAATGCTCAAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1800  
 DB 1741 AGAATGCTCAAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1800  
 QY 1801 CTATTCAAGATGAGAAAAATCTTAACCAATTAATTAATTAATTAATTAATTAATTA 1860  
 DB 1801 CTATTCAAGATGAGAAAAATCTTAACCAATTAATTAATTAATTAATTAATTAATTA 1860

QY	1861	AAAGCAAAAGCTGACCTTATAGCAGCAGATATAGCAATATATGGAATCGAATTTAAATATCC	1920
Db	1861	AAACCAAAAGCTGAACTTATAGCAGCAGATATATGCAATATGGAATCGAATTTAAATATCC	1920
QY	1921	ACAATTCAAAAGCAGCTTAAAAAGATAGCTGAGGAGAAAGCTTCTCCAGCGCATATTC	1980
Db	1921	ACAATTCAAAAGCAGCTTAAAAAGATAGCTGAGGAGAGAGCTTCTTACCGAGCATATTC	1980
QY	1981	ATGCGCTTGAACCTAGTACATAGCAAACTAAGGCCACCTAATTTGTACTCAATTTGCAGA	2040
Db	1981	ATGCGCTTGAACCTAGTACATAGCAAACTAAGGCCACCTAATTTGTACTCAATTTGCAGA	2040
QY	2041	TTGATAGTGTGTTCTAGCAGTGAGAGATTAAGAAAAAAAGTACACCAATGCCAGTCA	2100
Db	2041	TTGATAGTGTGTTCTAGCAGTGAGAGATTAAGAAAAAAAGTACACCAATGCCAGTCA	2100
QY	2101	GGCAGCAGAGAAACCTTACACCTCATGAGAGGTAAGAACTGCAACTGAGCCCAAGAGA	2160
Db	2101	GGCAGCAGAGAAACCTTACACCTCATGAGAGGTAAGAACTGCAACTGAGCCCAAGAGA	2160
QY	2161	GTAAACAGCCAAATGAGAGAGCAAGTAAGAAACATGACAGTGAATCTTTCCAGAGCTGA	2220
Db	2161	GTAAACAGCCAAATGAGAGAGCAAGTAAGAAACATGACAGTGAATCTTTCCAGAGCTGA	2220
QY	2221	AGTTAACAATAGCAGCTGCTCTTTACTAGTGTTCAAATACAGAGCAACTTAAAGAT	2280
Db	2221	AGTTAACAATAGCAGCTGCTCTTTACTAGTGTTCAAATACAGAGCAACTTAAAGAT	2280
QY	2281	TTGTCAATCTTAGCCTTCCAGAGAGAGAAAAAGAGAAACTAGAAACATTTAAAGGT	2340
Db	2281	TTGTCAATCTTAGCCTTCCAGAGAGAGAAAAAGAGAAACTAGAAACATTTAAAGGT	2340
QY	2341	CTAATTAATGCTGAGAGACCCCAAGATCTCATGTTAAGTGGAGAAAGGTTTTGCCAACTG	2400
Db	2341	CTAATTAATGCTGAGAGACCCCAAGATCTCATGTTAAGTGGAGAAAGGTTTTGCCAACTG	2400
QY	2401	AAAGATCTGTAGAGAGTAGCAGATATTTCAATGTCCTGCTAGTACGATTAATGAGCAGTG	2460
Db	2401	AAAGATCTGTAGAGAGTAGCAGATATTTCAATGTCCTGCTAGTACGATTAATGAGCAGTG	2460
QY	2461	AAAGATCTCTGTTAGTGAAGTTAGCAGCTAGGAGAGGCAAAAAACAGAACCAATTAAT	2520
Db	2461	AAAGATCTCTGTTAGTGAAGTTAGCAGCTAGGAGAGGCAAAAAACAGAACCAATTAAT	2520
QY	2521	GTGTGAGTCAAGTGCAGCATTTTGAACCCCAAGGCACTAATTCATGCTGTTCCAAAG	2580
Db	2521	GTGTGAGTCAAGTGCAGCATTTTGAACCCCAAGGCACTAATTCATGCTGTTCCAAAG	2580
QY	2581	ATAATAGAAATGACACAGAGGCTTTAAGTATCCATTGGACATGAAGTTAACCAAGTC	2640
Db	2581	ATAATAGAAATGACACAGAGGCTTTAAGTATCCATTGGACATGAAGTTAACCAAGTC	2640
QY	2641	GGGAACAACAGATAGAAATGGAAGAACTTGATGCTCAGTATTTTGCAGAAATCAT	2700
Db	2641	GGGAACAACAGATAGAAATGGAAGAACTTGATGCTCAGTATTTTGCAGAAATCAT	2700
QY	2701	TCAGAGTTTCAAAAGGCGCAGTCAATTTGCTGTTTTCAAATCCAGAGAAATGCCAAGAGG	2760
Db	2701	TCAGAGTTTCAAAAGGCGCAGTCAATTTGCTGTTTTCAAATCCAGAGAAATGCCAAGAGG	2760
QY	2761	AATGTCAGCAATCTCTCCCACTCTGGGCTCTTAAAGAAACAAAGTCCAAAGTCACTT	2820
Db	2761	AATGTCAGCAATCTCTCTCCCACTCTGGGCTCTTAAAGAAACAAAGTCCAAAGTCACTT	2820
QY	2821	TTTGAATGTGAACAAAAGAGAGAAATTCAGAGAAAGATGAGTCTAATATCAAGCTGTAC	2880
Db	2821	TTTGAATGTGAACAAAAGAGAGAAATTCAGAGAAAGATGAGTCTAATATCAAGCTGTAC	2880
QY	2881	AGACAGTTAATATCACTGACAGGCTTCTCTGTGTTGGTGCAGAAAGATTAAGCAGTTGATA	2940
Db	2881	AGACAGTTAATATCACTGACAGGCTTCTCTGTGTTGGTGCAGAAAGATTAAGCAGTTGATA	2940
QY	2941	ATGCCAAATGTAGTATCAAAAGAGGCTCTAGTGTGTCTATATCATCAAGTTCAGAGCA	3000
Db	2941	ATGCCAAATGTAGTATCAAAAGAGAGGCTCTAGTGTGTCTATATCATCAAGTTCAGAGCA	3000
QY	3001	ACGAACCTGAGACTCTTACTCCAAATTAACATGAGCTTTTACAAAACCAATATCTATAC	3060
Db	3001	ACGAACCTGAGACTCTTACTCCAAATTAACATGAGCTTTTACAAAACCAATATCTATAC	3060
QY	3061	CACACCTTTTCCATCAAGTCAATTTGTTTAAACATAATGTAAGAAAAATCTGCTAGAG	3120
Db	3061	CACACCTTTTCCATCAAGTCAATTTGTTTAAACATAATGTAAGAAAAATCTGCTAGAG	3120
QY	3121	AAAACCTTGAGAGAAATTCATATGTCACCTGGAAGAGAAATGGAATGAGAACATTTCCA	3180
Db	3121	AAAACCTTGAGAGAAATTCATATGTCACCTGGAAGAGAAATGGAATGAGAACATTTCCA	3180
QY	3181	GTACAGTGAAGCAATTAAGCCTAATAATTAAGAAATGTTTTTAAAGAACCCAGCT	3240
Db	3181	GTACAGTGAAGCAATTAAGCCTAATAATTAAGAAATGTTTTTAAAGAACCCAGCT	3240
QY	3241	CAGCAATATTAATAGTAGAGTTTCCAGTACTAATTAAGAGTGGGCTCCAGTATTAAGAA	3300
Db	3241	CAGCAATATTAATAGTAGAGTTTCCAGTACTAATTAAGAGTGGGCTCCAGTATTAAGAA	3300
QY	3301	TAGGTTCCAGTATGAAGAAACATTAAGAGAGAGTGAAGTGAAGAGAGGCTCCAGT	3360
Db	3301	TAGGTTCCAGTATGAAGAAACATTAAGAGAGAGTGAAGTGAAGAGAGGCTCCAGT	3360
QY	3361	ATGCTATGCTTGAATTAAGAGGTTTTTGCACACTGAGGCTTTAAACAAAGCTTCCGGA	3420
Db	3361	ATGCTATGCTTGAATTAAGAGGTTTTTGCACACTGAGGCTTTAAACAAAGCTTCCGGA	3420
QY	3421	GTATATTGTAGCATCTCTAATTAAGAAAGCAAGATATGAAGAGTATGTTACAGCTGTTA	3480
Db	3421	GTATATTGTAGCATCTCTAATTAAGAAAGCAAGATATGAAGAGTATGTTACAGCTGTTA	3480
QY	3481	ATACAGATTTCTCTCATATCTGATTTCAAGTAACTTAAGAAAGCCTATGAGAGAGTGC	3540
Db	3481	ATACAGATTTCTCTCATATCTGATTTCAAGTAACTTAAGAAAGCCTATGAGAGAGTGC	3540
QY	3541	ATGCATCTCAGGTTTGTCTGAGACACCTGATGACCTGATGATGATGATGAATTAAGG	3600
Db	3541	ATGCATCTCAGGTTTGTCTGAGACACCTGATGACCTGATGATGATGATGAATTAAGG	3600
QY	3601	AAAGATCTAGATTTGCTGAAATTAAGCAATTAAGAAAGTCTGCTGTTTACCAAAAGG	3660
Db	3601	AAAGATCTAGATTTGCTGAAATTAAGCAATTAAGAAAGTCTGCTGTTTACCAAAAGG	3660
QY	3661	TCACAGAAAGAGAGCTTATGAGAGAGTCTAGCCCTTACACCAATTAACATTTGCTCAGG	3720
Db	3661	TCACAGAAAGAGAGCTTATGAGAGAGTCTAGCCCTTACACCAATTAACATTTGCTCAGG	3720
QY	3721	GTTTACCAGAGAGGCGCAGAAATTAAGTCTCAGAGAGAACTTATCTAGTGAAGT	3780
Db	3721	GTTTACCAGAGAGGCGCAGAA	

|||||  
 Db 4021 GTGATTGGAGACTCTTACGCAATACAAACACCCGAGATCCCTTTGTTGATGTTCTT 4080  
 OY 4081 CCAACAAATGAGGCATCAGTCTGAAAGCCAGGGAGTTGGTCTGAGTGACAGGAATGG 4140  
 Db 4081 CCAACAAATGAGGCATCAGTCTGAAAGCCAGGGAGTTGGTCTGAGTGACAGGAATGG 4140  
 OY 4141 TTTCAGTGTATGAAGAAAGAGAAAGGGCTTGGAGAAATTAATCAAGAGAGCA 4200  
 Db 4141 TTTCAGTGTATGAAGAAAGAGAAAGGGCTTGGAGAAATTAATCAAGAGAGCA 4200  
 OY 4201 TGGATTCAACTTAGTGAAGACACATCTGGTGTGAGAGTGAAGCAAGGCTCTGAG 4260  
 Db 4201 TGGATTCAACTTAGTGAAGACACATCTGGTGTGAGAGTGAAGCAAGGCTCTGAG 4260  
 OY 4261 ACTGCTAGGAGCTTCTCTCAGAGTACATTTAACTTACCTCAGCAGAGGATACCATGC 4320  
 Db 4261 ACTGCTAGGAGCTTCTCTCAGAGTACATTTAACTTACCTCAGCAGAGGATACCATGC 4320  
 OY 4321 AACATACTGATTAAGCTCAGCAGAGAAATGCTGAAGTGAAGTGTGTAGAACAGC 4380  
 Db 4321 AACATACTGATTAAGCTCAGCAGAGAAATGCTGAAGTGAAGTGTGTAGAACAGC 4380  
 OY 4381 ATGGAGCCAGCCTTCTPAACAGCTACCCCTTCATCATPAAGTACTCTTCCCTTGAAG 4440  
 Db 4381 ATGGAGCCAGCCTTCTPAACAGCTACCCCTTCATCATPAAGTACTCTTCCCTTGAAG 4440  
 OY 4441 ACCTGGCAATCCGAACAAAGCAGATCAGAAAGAGTATTAAGTCTTCACAGAAAGTA 4500  
 Db 4441 ACCTGGCAATCCGAACAAAGCAGATCAGAAAGAGTATTAAGTCTTCACAGAAAGTA 4500  
 OY 4501 GTGAATACCCCTATTAAGCCAGAAATCCAGAGGCTTCTGCTGACAGTTTGAAGTGTCTG 4560  
 Db 4501 GTGAATACCCCTATTAAGCCAGAAATCCAGAGGCTTCTGCTGACAGTTTGAAGTGTCTG 4560  
 OY 4561 CAGATAGTCTTACCAAGTAAATTAAGAACCCAGAGAGTGAAGAGTCAATCCCTTCTAAT 4620  
 Db 4561 CAGATAGTCTTACCAAGTAAATTAAGAACCCAGAGAGTGAAGAGTCAATCCCTTCTAAT 4620  
 OY 4621 GCCCATATTAGATGATGGGTGACATGACAGTCTGCTGAGAGCTTCCAGATAGAA 4680  
 Db 4621 GCCCATATTAGATGATGGGTGACATGACAGTCTGCTGAGAGCTTCCAGATAGAA 4680  
 OY 4681 ACTACCATCTCAAGAGAGCTCATTAAGTGTGTGATGAGAGCAACAGCTGGAAG 4740  
 Db 4681 ACTACCATCTCAAGAGAGCTCATTAAGTGTGTGATGAGAGCAACAGCTGGAAG 4740  
 OY 4741 AGTCTGGGCCACAGATTTGACGAAACATCTTACTTGCCAGGCAAGATCTAGAGGAA 4800  
 Db 4741 AGTCTGGGCCACAGATTTGACGAAACATCTTACTTGCCAGGCAAGATCTAGAGGAA 4800  
 OY 4801 CCCCTTACCTGSAATCTGSAATCAGGCTTCTCTGATGACCCCTGATCTGATCTCTG 4860  
 Db 4801 CCCCTTACCTGSAATCTGSAATCAGGCTTCTCTGATGACCCCTGATCTGATCTCTG 4860  
 OY 4861 AAGACAGAGCCAGAGTCTGCTGTGGCAACATACCATCTTCAACTCTGCAATGA 4920  
 Db 4861 AAGACAGAGCCAGAGTCTGCTGTGGCAACATACCATCTTCAACTCTGCAATGA 4920  
 OY 4921 AAGTCCCAATTTGAAGTTGAGATCTGCCAGAGTCCAGGCTGCTCATCTACTG 4980  
 Db 4921 AAGTCCCAATTTGAAGTTGAGATCTGCCAGAGTCCAGGCTGCTCATCTACTG 4980  
 OY 4981 ATACTGCTGGGTATTAATGCAATGGAAGAGTGTGACAGAGGAAACCCGAATTTGACAG 5040  
 Db 4981 ATACTGCTGGGTATTAATGCAATGGAAGAGTGTGACAGAGGAAACCCGAATTTGACAG 5040  
 OY 5041 CTTCAACAGAAAGGCTCAACAAAAGATGTCAATGTGTGTGTGCTGACCCAGAG 5100  
 Db 5041 CTTCAACAGAAAGGCTCAACAAAAGATGTCAATGTGTGTGTGCTGACCCAGAG 5100  
 OY 5101 AATTATGCTGCTGTACAGTTTGCAGAAACACCATCTTAACTTAATCTAATTA 5160  
 Db 5101 AATTATGCTGCTGTGTACAGTTTGCAGAAACACCATCTTAACTTAATCTAATTA 5160

Db 5101 AATTATGCTGCTGTGTACAGTTTGCAGAAACACCATCTTAACTTAATCTAATTA 5160  
 OY 5161 CTGAGAGACTACTCATGTTTGTATGAAACAGATGCTGATGTTGTGTGACAGCAGC 5220  
 Db 5161 CTGAGAGACTACTCATGTTTGTATGAAACAGATGCTGATGTTGTGTGACAGCAGC 5220  
 OY 5221 TGAATATTTTCTAGGAATTCGGGAGAGAAATGGGTAGTTAGCTATTTCTGGGTGACCC 5280  
 Db 5221 TGAATATTTTCTAGGAATTCGGGAGAGAAATGGGTAGTTAGCTATTTCTGGGTGACCC 5280  
 OY 5281 AGTCTATTAAGAAAGAAATGCTGAATGACATGATTTTGAAGTCAAGAGATGTGG 5340  
 Db 5281 AGTCTATTAAGAAAGAAATGCTGAATGACATGATTTTGAAGTCAAGAGATGTGG 5340  
 OY 5341 TCAATGGAAGAAACCAAGGCTCCAAAGCAGCAAGAAATCCAGAGCAAGAAATCT 5400  
 Db 5341 TCAATGGAAGAAACCAAGGCTCCAAAGCAGCAAGAAATCCAGAGCAAGAAATCT 5400  
 OY 5401 TCAGGGGGCTAGAAATCTGTTGCTATGAGGCTTCCACCAACATGCCACAGATCACTGG 5460  
 Db 5401 TCAGGGGGCTAGAAATCTGTTGCTATGAGGCTTCCACCAACATGCCACAGATCACTGG 5460  
 OY 5461 AATGATGTGACAGCTGTGTGTCTCTGCTGTGTGAAGAGCTTTCATCATCCCTTG 5520  
 Db 5461 AATGATGTGACAGCTGTGTGTCTCTGCTGTGTGAAGAGCTTTCATCATCCCTTG 5520  
 OY 5521 GCACAGGTGTCACCAATGTTGTTGTGACAGCCAGATGCTGACAGAGAGCAATGGCT 5580  
 Db 5521 GCACAGGTGTCACCAATGTTGTTGTGACAGCCAGATGCTGACAGAGAGCAATGGCT 5580  
 OY 5581 TCCATGCAATTTGGGACAGATGTGTGAGGACCTGTGTGTGACCCGAGAGTGGTGTGACA 5640  
 Db 5581 TCCATGCAATTTGGGACAGATGTGTGAGGACCTGTGTGTGACCCGAGAGTGGTGTGACA 5640  
 OY 5641 GTTAGCAGCTTACCAAGTCCAGAGGCTGGACACCTTACCTGATACCCAGATCCCCACA 5700  
 Db 5641 GTTAGCAGCTTACCAAGTCCAGAGGCTGGACACCTTACCTGATACCCAGATCCCCACA 5700  
 OY 5701 GCCCAGTACTGA 5711  
 Db 5701 GCCCAGTACTGA 5711

RESULT 10  
 AAY46461  
 ID AAY46461 standard; cDNA: 5711 BP.  
 XX  
 AC AAY46461;  
 XX  
 DT 18-NOV-1998 (first entry)  
 XX  
 DE Human BRCA1 omi2 polymorphism #4 cDNA.  
 KW BRCA1; omi2; human; breast and ovarian cancer predisposing gene;  
 KW polymorphism; susceptibility; anti-oncogene; tumour suppressor;  
 KW chromosome 17q; ss.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 120..5711  
 FT FT /\*tag= a  
 FT FT /product= "BRCA1 omi2 protein"  
 FT variation 3232  
 FT FT /\*tag= b  
 FT FT /note= "This polymorphic variation can be an A or G  
 PN US5750400-A.  
 XX  
 PD 12-MAY-1998.  
 XX  
 PF 12-FEB-1997; 97US-0798691.







QY 1621 AATTAAAGCTAAAGAGACCTACATCGCCCTTCATCTGAGATTTTATCAAGAAAG 1680  
DB 1621 AATTAAAGCTAAAGAGACCTACATCGCCCTTCATCTGAGATTTTATCAAGAAAG 1680  
QY 1681 CAGATTTGGAGCTCAAAAGACCTCCTGAATGATTAATTCAGGAGACTRACCAAGCGAGC 1740  
DB 1681 CAGATTTGGAGCTCAAAAGACCTCCTGAATGATTAATTCAGGAGACTRACCAAGCGAGC 1740  
QY 1741 AGAATGCTCAAGTGAATTAATTTACTAATAGTGTCTAGAAATTAATAACAAAGGTATT 1800  
DB 1741 AGAATGCTCAAGTGAATTAATTTACTAATAGTGTCTAGAAATTAATAACAAAGGTATT 1800  
QY 1801 CTATTCAGAAATGAGAAAAATCTTAACCAATAGAAATCACTCGAAAAAGAAATCTGCTTCA 1860  
DB 1801 CTATTCAGAAATGAGAAAAATCTTAACCAATAGAAATCACTCGAAAAAGAAATCTGCTTCA 1860  
QY 1861 AAACGAAGCTGAACCTTAAGAGAGAGCTATTAAGCATATGGAATCTGAAATTAATATCC 1920  
DB 1861 AAACGAAGCTGAACCTTAAGAGAGAGCTATTAAGCATATGGAATCTGAAATTAATATCC 1920  
QY 1921 ACAATTCAAAAGACACCTTAAGAAATAGGCTGAGAGAGAGAGCTTCTACAGCATATTC 1980  
DB 1921 ACAATTCAAAAGACACCTTAAGAAATAGGCTGAGAGAGAGAGCTTCTACAGCATATTC 1980  
QY 1981 ATGCGCTTGAACCTAGTAGTGAATCTTAAGCCCACTAATTTGACTGAAATTTGCAAA 2040  
DB 1981 ATGCGCTTGAACCTAGTAGTGAATCTTAAGCCCACTAATTTGACTGAAATTTGCAAA 2040  
QY 2041 TTGATATTTGTTCTAGAGAGTGAAGATTAAGAAAAAAGTACAACCAATGCGACTCA 2100  
DB 2041 TTGATATTTGTTCTAGAGAGTGAAGATTAAGAAAAAAGTACAACCAATGCGACTCA 2100  
QY 2101 GGCACACAGAAACCTCACTCATGGAAGTAAAGAACCTGCACTGAGACCAAGAGA 2160  
DB 2101 GGCACACAGAAACCTCACTCATGGAAGTAAAGAACCTGCACTGAGACCAAGAGA 2160  
QY 2161 GTAAACAAGCCAAATGAACAGACAGATTAAGAACATGACAGTACTTTCCACAGAGCTGA 2220  
DB 2161 GTAAACAAGCCAAATGAACAGACAGATTAAGAACATGACAGTACTTTCCACAGAGCTGA 2220  
QY 2221 AGTTAAACAATGCACTCGGTTCTTTTACTAAGTGTCAATATACAGTGAATTAAGAAAT 2280  
DB 2221 AGTTAAACAATGCACTCGGTTCTTTTACTAAGTGTCAATATACAGTGAATTAAGAAAT 2280  
QY 2281 TTGTCATCTAGACCTTCCAAAGAGAAAAAGAAAAAGAAACTGAAACACTTAAAGTGT 2340  
DB 2281 TTGTCATCTAGACCTTCCAAAGAGAAAAAGAAAAAGAAACTGAAACACTTAAAGTGT 2340  
QY 2341 CTAAATAATGCTGAAGACCCCAAGATCTCATGTTAAGTGAAGAAAGGTTTGGCAACTG 2400  
DB 2341 CTAAATAATGCTGAAGACCCCAAGATCTCATGTTAAGTGAAGAAAGGTTTGGCAACTG 2400  
QY 2401 AAAGATCTGAGAGAGTGAAGTATTTCAATGCTGACTGACTGATTAAGGCACTGAGC 2460  
DB 2401 AAAGATCTGAGAGAGTGAAGTATTTCAATGCTGACTGACTGATTAAGGCACTGAGC 2460  
QY 2461 AAAGATCTGCTTACTGGAAGTGAAGTGAAGTGAAGGCAAAAAAGCAAGCAATTAAT 2520  
DB 2461 AAAGATCTGCTTACTGGAAGTGAAGTGAAGTGAAGGCAAAAAAGCAAGCAATTAAT 2520  
QY 2521 GTGTCAGTCAAGTGTGACAGATTTGAAAAACCCCAAGGAGCACTAATTCATGTTGCCAAG 2580  
DB 2521 GTGTCAGTCAAGTGTGACAGATTTGAAAAACCCCAAGGAGCACTAATTCATGTTGCCAAG 2580  
QY 2581 ATATATAAATGACACAGAGAGCTTTAAGTATCCATTGGGACACAGAAAGTTAACCAAGTC 2640  
DB 2581 ATATATAAATGACACAGAGAGCTTTAAGTATCCATTGGGACACAGAAAGTTAACCAAGTC 2640  
QY 2641 GGGAAACAGCATGGAATGGAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2700  
DB 2641 GGGAAACAGCATGGAATGGAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2700  
QY 2701 TCAAGTTTCCAAAGCGGCACTGATTTGCTGTTTCAAAATCCAGAAATGCAAGAGAGG 2760

DB 2701 TCAAGTTTCCAAAGCGGCACTGATTTGCTGTTTCAAAATCCAGAAATGCAAGAGAGG 2760  
QY 2761 AATGCAACATTTCTGCGCCACCTGCGGCTTAAAGCAAAAGTCCAAAGCACTT 2820  
DB 2761 AATGCAACATTTCTGCGCCACCTGCGGCTTAAAGCAAAAGTCCAAAGCACTT 2820  
QY 2821 TTGAATGGAACAAAGAGAGAAAAATCAAGAAAGAAATGAGTCTAATTCAGAGCTGTAC 2880  
DB 2821 TTGAATGGAACAAAGAGAGAAAAATCAAGAAAGAAATGAGTCTAATTCAGAGCTGTAC 2880  
QY 2881 AGACAGTTAATATCACTGACAGGCTTTCTGTTGTTGTCAGAAAAGATTAAGCAGTTGATA 2940  
DB 2881 AGACAGTTAATATCACTGACAGGCTTTCTGTTGTTGTCAGAAAAGATTAAGCAGTTGATA 2940  
QY 2941 ATGCGAAATGATAGTATCAAGAGAGGCTAGGTTTGTCTATCATCTCAGTTCAAGAGCA 3000  
DB 2941 ATGCGAAATGATAGTATCAAGAGAGGCTAGGTTTGTCTATCATCTCAGTTCAAGAGCA 3000  
QY 3001 ACAGAACTGACATTAATCAATTAACATGACCTTTTACAAAACCCATATGCTATAC 3060  
DB 3001 ACAGAACTGACATTAATCAATTAACATGACCTTTTACAAAACCCATATGCTATAC 3060  
QY 3061 CACCACCTTTTCCATCAAGTCAATTTGTTAAACTAAATGTAAGAAAAATGCTGACAGG 3120  
DB 3061 CACCACCTTTTCCATCAAGTCAATTTGTTAAACTAAATGTAAGAAAAATGCTGACAGG 3120  
QY 3121 AAAACTTTGAGAGACATTTCAATGTCACCTGAAAGAAAGAAATGGAACAAATTCOA 3180  
DB 3121 AAAACTTTGAGAGACATTTCAATGTCACCTGAAAGAAAGAAATGGAACAAATTCOA 3180  
QY 3181 GTACAGTGCAGACATTAAGCCGTAAATTAAGAGAAATGTTTAAAGAACCCAGCT 3240  
DB 3181 GTACAGTGCAGACATTAAGCCGTAAATTAAGAGAAATGTTTAAAGAACCCAGCT 3240  
QY 3241 CAAGCAATTAATTAAGAGTGAAGTTCACAGTAAATGAAGTGGCTCCAGTAAATGA 3300  
DB 3241 CAAGCAATTAATTAAGAGTGAAGTTCACAGTAAATGAAGTGGCTCCAGTAAATGA 3300  
QY 3301 TAGGTTCCAGTGAAGAAACATCAAGAGAGAAAGTGAAGAGAGGCGCAAAATTTGA 3360  
DB 3301 TAGGTTCCAGTGAAGAAACATCAAGAGAGAAAGTGAAGAGAGGCGCAAAATTTGA 3360  
QY 3361 ATGCTATGCTTGAATGAGGCTTTTGGCACTGAGAGCTATTAACCAAGTCTCTCGAA 3420  
DB 3361 ATGCTATGCTTGAATGAGGCTTTTGGCACTGAGAGCTATTAACCAAGTCTCTCGAA 3420  
QY 3421 GTAAATTTGAGCATCTGGAATTAAGAAAGCAAGATTAAGAAAGTGAAGTGAAGTGA 3480  
DB 3421 GTAAATTTGAGCATCTGGAATTAAGAAAGCAAGATTAAGAAAGTGAAGTGAAGTGA 3480  
QY 3481 ATACAGATTTCTCCATATCTGATGATTAAGTGAAGTGAAGTGAAGTGAAGTGA 3540  
DB 3481 ATACAGATTTCTCCATATCTGATGATTAAGTGAAGTGAAGTGAAGTGAAGTGA 3540  
QY 3541 ATGCAATCTCAGGTTTGTCTGAGACACCTGATGACCTGTTAGATGATGTTGAATTAAGG 3600  
DB 3541 ATGCAATCTCAGGTTTGTCTGAGACACCTGATGACCTGTTAGATGATGTTGAATTAAGG 3600  
QY 3601 AAGTACTAGTTTGTCTGAGAAATGACATTAAGAAAGTGTGCTGTTTAAAGCAAGCG 3660  
DB 3601 AAGTACTAGTTTGTCTGAGAAATGACATTAAGAAAGTGTGCTGTTTAAAGCAAGCG 3660  
QY 3661 TCCAGAAAGAGAGCTTGAAGAGAGTCTTACCCCTTACCCATACACATTTGCTCAGG 3720  
DB 3661 TCCAGAAAGAGAGCTTGAAGAGAGTCTTACCCCTTACCCATACACATTTGCTCAGG 3720  
QY 3721 GTTACCGAAGAGGCGCAAGAAATTAAGTCTCAGAGAAACCTTATCTGAGAGATG 3780  
DB 3721 GTTACCGAAGAGGCGCAAGAAATTAAGTCTCAGAGAAACCTTATCTGAGAGATG 3780  
QY 3781 AAGAGCTTCCCTGCTTCCAAACATTTGTTATTTGTTAAAGTAAACATATACCTTCTAGT 3840

Db	3781	AAGAGCTCCCTCCCTCCAACTTGTATTGTTAAAGTAAACAATATACCTTCTAGT	3840
Oy	3841	CTACTAGGATAGACACGCTGTGTACGAGTGTCTGTCTAAGACACAGAGAGAAATTAAT	3900
Db	3841	CTACTAGGATAGACACGCTGTGTACGAGTGTCTGTCTAAGACACAGAGAGAAATTAAT	3900
Oy	3901	TATCATTTGAAGATAGCTTAAATGACTGCAGTAACCAAGTAAATTTGGCAAGGCATCTC	3960
Db	3901	TATCATTTGAAGATAGCTTAAATGACTGCAGTAACCAAGTAAATTTGGCAAGGCATCTC	3960
Oy	3961	AGGACATACCTTAGTGAAGAAACAATATGTTCTGTAGCTTGTCTTCTTACAGTGA	4020
Db	3961	AGGACATACCTTAGTGAAGAAACAATATGTTCTGTAGCTTGTCTTCTTACAGTGA	4020
Oy	4021	GTGAATTTGAAGATGACTGCAGTAACCAACCCAGATCCCTTCTTGTAGTGGTCTT	4080
Db	4021	GTGAATTTGAAGATGACTGCAGTAACCAACCCAGATCCCTTCTTGTAGTGGTCTT	4080
Oy	4081	CCAAACAAATGAGCATCTGAAAGCCAGGAGTGGTCTGAGTGAACAGGAATTTG	4140
Db	4081	CCAAACAAATGAGCATCTGAAAGCCAGGAGTGGTCTGAGTGAACAGGAATTTG	4140
Oy	4141	TTTTCAGATGATGAAGAAAGAGAGCGGCTTTGGAGAAATTAATCAAGAGAGCAAGCA	4200
Db	4141	TTTTCAGATGATGAAGAAAGAGAGCGGCTTTGGAGAAATTAATCAAGAGAGCAAGCA	4200
Oy	4201	TGAGATTCAACTTAGTGAAGAGAGCATCTGGTGTGAGTGAAGCAAGCGCTCTGAG	4260
Db	4201	TGAGATTCAACTTAGTGAAGAGAGCATCTGGTGTGAGTGAAGCAAGCGCTCTGAG	4260
Oy	4261	ACTGCTCAGGGCTATCTCTCAGAGTGTACATTTTAACTCAGCAGAGGATACATGC	4320
Db	4261	ACTGCTCAGGGCTATCTCTCAGAGTGTACATTTTAACTCAGCAGAGGATACATGC	4320
Oy	4321	AACATTAACCTGATAAGCTCCACAGAGAAATGGCTGAACATAGAACTGTTGAAGACG	4380
Db	4321	AACATTAACCTGATAAGCTCCACAGAGAAATGGCTGAACATAGAACTGTTGAAGACG	4380
Oy	4381	ATGGAGGAGGCTTCTTACAGCTACCTTCATCAATGAAGTCTTCTGCCCCCTTGAAG	4440
Db	4381	ATGGAGGAGGCTTCTTACAGCTACCTTCATCAATGAAGTCTTCTGCCCCCTTGAAG	4440
Oy	4441	ACCTGGGAATCCAGAACAAAGCAGATCAAGAAAAGCAGTATTAACTTCACAGAAAAGTA	4500
Db	4441	ACCTGGGAATCCAGAACAAAGCAGATCAAGAAAAGCAGTATTAACTTCACAGAAAAGTA	4500
Oy	4501	GTGAATACCTTATTAACCAAGATCCAGAGGCTTCTGTCGAACAATTTGAGGTGCTG	4560
Db	4501	GTGAATACCTTATTAACCAAGATCCAGAGGCTTCTGTCGAACAATTTGAGGTGCTG	4560
Oy	4561	CAGATAGTTCTACAGTAAATAAAGAACAGAGTGAAGGTCATCCCTTCTTAAT	4620
Db	4561	CAGATAGTTCTACAGTAAATAAAGAACAGAGTGAAGGTCATCCCTTCTTAAT	4620
Oy	4621	GGCCATCATTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	4680
Db	4621	GGCCATCATTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	4680
Oy	4681	ACTACCATCTCAAGAGAGGATCATTAAGTGTGTGATGATGATGATGATGATGATGATG	4740
Db	4681	ACTACCATCTCAAGAGAGGATCATTAAGTGTGTGATGATGATGATGATGATGATGATG	4740
Oy	4741	AGTCTGGGACACAGATTTGACGGAACATCTTACTTGGCCAGGCAAGATCTAGAGGAA	4800
Db	4741	AGTCTGGGACACAGATTTGACGGAACATCTTACTTGGCCAGGCAAGATCTAGAGGAA	4800
Oy	4801	CCCCTTACCTGGAATCTGAATCAGCTCTTCTCTGATGACCTGATCTGATCTCTG	4860
Db	4801	CCCCTTACCTGGAATCTGAATCAGCTCTTCTCTGATGACCTGATCTGATCTCTG	4860
Oy	4861	AAGACAGAGCCCAAGAGTCAAGTCTGTTGGCAACATTTCAACTCTGACATTTGA	4920
Db	4861	AAGACAGAGCCCAAGAGTCAAGTCTGTTGGCAACATTTCAACTCTGACATTTGA	4920
Oy	4921	AAGTCCCAATTGAAAGTTGCAGAAATCTGCCAGAGTCCAGCTGCTCATACTAGC	4980
Db	4921	AAGTCCCAATTGAAAGTTGCAGAAATCTGCCAGAGTCCAGCTGCTCATACTAGC	4980
Oy	4981	ATACTGCTGGGTATTAATGCAATGGAAGAAAGTGTGACAGGAGAGACCAATTTGACAG	5040
Db	4981	ATACTGCTGGGTATTAATGCAATGGAAGAAAGTGTGACAGGAGAGACCAATTTGACAG	5040
Oy	5041	CTTCAACAGAAAGGTCACAAAGAAATGCTTGGTGTGCTGCTGCTGCTGCTGCTGCTG	5100
Db	5041	CTTCAACAGAAAGGTCACAAAGAAATGCTTGGTGTGCTGCTGCTGCTGCTGCTGCTG	5100
Oy	5101	AATTTATGCTGCTGTCAGATTTGCAAGAAATCCACATCACTTAATTAATTAATTA	5160
Db	5101	AATTTATGCTGCTGTCAGATTTGCAAGAAATCCACATCACTTAATTAATTAATTA	5160
Oy	5161	CTGAGAGACTACTCATGTTGTTATGAAACACATGCTGAGTTGTGTGTAAGCGACAC	5220
Db	5161	CTGAGAGACTACTCATGTTGTTATGAAACACATGCTGAGTTGTGTGTAAGCGACAC	5220
Oy	5221	TGAATATTTTCTAGGAATTTGGCGGAGAAATGGGTACTTACATTTCTGGTGACCC	5280
Db	5221	TGAATATTTTCTAGGAATTTGGCGGAGAAATGGGTACTTACATTTCTGGTGACCC	5280
Oy	5281	AGTCTATTAAAGAAAGAAATGCTGAATGAGCATGATTTTGAAGTCAAGAGAGATGAG	5340
Db	5281	AGTCTATTAAAGAAAGAAATGCTGAATGAGCATGATTTTGAAGTCAAGAGAGATGAG	5340
Oy	5341	TCAATGGAAGAAACCAAGAGTCCAAAGGAGCAAGAGATCCAGACAGAAAGATCT	5400
Db	5341	TCAATGGAAGAAACCAAGAGTCCAAAGGAGCAAGAGATCCAGACAGAAAGATCT	5400
Oy	5401	TCAAGGGGCTAGAAATCTGTTGCTATGAGGCTTCCCAACATGCCCACAGATCAACTGG	5460
Db	5401	TCAAGGGGCTAGAAATCTGTTGCTATGAGGCTTCCCAACATGCCCACAGATCAACTGG	5460
Oy	5461	AATGATGATGACAGTGTGTGCTTCTGTGTGTAAGAGGCTTTCATCATTTACCCCTTG	5520
Db	5461	AATGATGATGACAGTGTGTGCTTCTGTGTGTAAGAGGCTTTCATCATTTACCCCTTG	5520
Oy	5521	GCACAGGTGTCACCAATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	5580
Db	5521	GCACAGGTGTCACCAATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	5580
Oy	5581	TCCATGCAATTTGGGACAGATGTGTGAGGCACTGTGTGTGTGTGTGTGTGTGTGTGTGT	5640
Db	5581	TCCATGCAATTTGGGACAGATGTGTGAGGCACTGTGTGTGTGTGTGTGTGTGTGTGTGT	5640
Oy	5641	GTGTAGACCTTACAGATGTCAGAGAGCTTGAGACACCTTACCTGATACCCAGATCCCCACA	5700
Db	5641	GTGTAGACCTTACAGATGTCAGAGAGCTTGAGACACCTTACCTGATACCCAGATCCCCACA	5700
Oy	5701	GGCCTTACCTGA 5711	
Db	5701	GGCCTTACCTGA 5711	

RESULT 11  
AAV46462...  
ID AAV46462 standard; cDNA; 5711 BP.  
XX  
AC AAV46462;  
DT 18-NOV-1998 (first entry)  
XX  
DE Human BRCA1 omi2 polymorphism #5 cDNA.  
XX  
KW BRCA1; omi2; human; breast and ovarian cancer predisposing gene;  
KW polymorphism; susceptibility; anti-oncogene; tumour suppressor;  
KW chromosome 17q; ss.  
XX  
OS Homo sapiens.

XX Key Location/Qualifiers  
 FH CDS 120..5711  
 FT /\*tag= a  
 FT /product= "BRCA1 omi2 protein"  
 FT variation 3667  
 FT /\*tag= b  
 FT /note= "this polymorphic variation can be an A or G nucleotide"

US5750400-A.  
 12-MAY-1998.  
 12-FEB-1997; 97US-0798691.  
 12-FEB-1996; 96US-0598591.  
 12-FEB-1997; 97US-0798691.  
 (ONCO-) ONCORMED INC.  
 Allen AC, Alvares CP, Critz BS, Murphy PD, Olson SJ;  
 Schelter DB, Zeng B;  
 WPI: 1998-296774/26.  
 BRCA1 omi gene coding sequences - useful for distinguishing between polymorphisms and mutation(s) in the screening for disposition to breast or ovarian cancer  
 Claim 2e; Page -: 54pp; English.

This sequence encodes a human BRCA1 (breast and ovarian cancer predisposing gene) omi2 gene in which a polymorphic variation occurs at nucleotide 3667. This sequence and other polymorphic variations of this sequence are useful for the identification of an individual who may or may not have an increased susceptibility to breast or ovarian cancer. The sequences used identify gene changes which are due to polymorphisms, rather than cancer causing mutations. BRCA1 is an anti-oncogene (tumour suppressor) which is involved in genetic inheritance of cancers, especially breast and ovarian cancer. It is found at human chromosome 17q which is known to be linked to cancer susceptibility, especially breast cancer. Cells containing a mutation in this gene lose the wild-type function of BRCA1 and are more susceptible to cancers.  
 NOTE: This sequence does not appear in the specification but has been created from the wild type BRCA1 omi2 gene represented in AAV46449.

Sequence 5711 BP: 1955 A; 1098 C; 1274 G; 1383 T; 1 other:

Query Match 100.0%; Score 5709; DB 19; Length 5711;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 5709; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 AGCTCGCTGAGACTTCTGAGACCCCGACAGGCTGTGGGTTTCTCAGATTAAGTGGCC 60  
 DB 1 AGCTCGCTGAGACTTCTGAGACCCCGACAGGCTGTGGGTTTCTCAGATTAAGTGGCC 60  
 OY 61 CCTCGGCTCAGAGAGCCCTTCACCTGCTGTGGTAAAGTTCAATGGAAAGAA 120  
 DB 61 CCTCGGCTCAGAGAGCCCTTCACCTGCTGTGGTAAAGTTCAATGGAAAGAA 120  
 OY 121 TGGATTATCTGCTTCTGCGCTGAGAGTAAAGTAAATGCTAATGCTATGAGAAA 180  
 DB 121 TGGATTATCTGCTTCTGCGCTGAGAGTAAAGTAAATGCTAATGCTATGAGAAA 180  
 OY 181 TCTTAGAGTGTCCATCTGTCTGAGTGTATCAAGGAACCTGTCTCACAAGAGTGAC 240  
 DB 181 TCTTAGAGTGTCCATCTGTCTGAGTGTATCAAGGAACCTGTCTCACAAGAGTGAC 240  
 OY 241 ACATATTTTGCATTTTGCATGCTGAACCTTCACCCAGAAAGAGGCTTCACAGT 300  
 DB 241 ACATATTTTGCATTTTGCATGCTGAACCTTCACCCAGAAAGAGGCTTCACAGT 300

OY 301 GTCTTTATGTAGAATGATATTAACCAAAAGAGCTTACAAAGAAAGTACGAGATTAGTC 360  
 DB 301 GTCTTTATGTAGAATGATATTAACCAAAAGAGCTTACAAAGAAAGTACGAGATTAGTC 360  
 OY 361 AACTTTGTAAGAGCTATTGAAAATCATTTGTGCTTTCAGCTTGACACAGTTTGAGCT 420  
 DB 361 AACTTTGTAAGAGCTATTGAAAATCATTTGTGCTTTCAGCTTGACACAGTTTGAGCT 420  
 OY 421 ATGCAAAAGCTATATTTTGCAAAAGAAAATTAACCTTCCTGATATATCAAAAGATG 480  
 DB 421 ATGCAAAAGCTATATTTTGCAAAAGAAAATTAACCTTCCTGATATATCAAAAGATG 480  
 OY 481 AAGTTTCTATCATCCAAAGTATGGCTACAGAAACCGTCCCAAAAGACTTCTACAGAGTG 540  
 DB 481 AAGTTTCTATCATCCAAAGTATGGCTACAGAAACCGTCCCAAAAGACTTCTACAGAGTG 540  
 OY 541 AACCAGAAATCTTCTTTCGAGAGAAACCACTGTCTGCTCAACTCTTAACCTTGGAA 600  
 DB 541 AACCAGAAATCTTCTTTCGAGAGAAACCACTGTCTGCTCAACTCTTAACCTTGGAA 600  
 OY 601 CTGTGAGAACTCTGAGACAAAGAGGATACAAACCTCAAAAGAGCTGTCTACATATG 660  
 DB 601 CTGTGAGAACTCTGAGACAAAGAGGATACAAACCTCAAAAGAGCTGTCTACATATG 660  
 OY 661 AATTGGATCTGATTCCTCTGAGAGATACCTTTAATTAAGCAACTATTGACAGTGTGGAG 720  
 DB 661 AATTGGATCTGATTCCTCTGAGAGATACCTTTAATTAAGCAACTATTGACAGTGTGGAG 720  
 OY 721 ATCAAGATTTGTACAAATCACCCCTCAAGAACAGGATGAATACGTTTGCATTTCTG 780  
 DB 721 ATCAAGATTTGTACAAATCACCCCTCAAGAACAGGATGAATACGTTTGCATTTCTG 780  
 OY 781 CAAAAGAGCTGCTGTGATTTCTGAGAGGATGTATACAAATTCGATACATCATCTAC 840  
 DB 781 CAAAAGAGCTGCTGTGATTTCTGAGAGGATGTATACAAATTCGATACATCATCTAC 840  
 OY 841 CCAGTAATATGATTTGACACCACTGAGAGCGTGCAGTGAAGCGCATCCAGAAAAGT 900  
 DB 841 CCAGTAATATGATTTGACACCACTGAGAGCGTGCAGTGAAGCGCATCCAGAAAAGT 900  
 OY 901 ATCAGGATAGTTCCTTCAAACTTGCATGTGAGACCATGTGCGCAAAATCTCATGCCA 960  
 DB 901 ATCAGGATAGTTCCTTCAAACTTGCATGTGAGACCATGTGCGCAAAATCTCATGCCA 960  
 OY 961 GCTATTACAGCAGAGAACGAGTATTAATACATCAAAAGACAGATGAATGAGAA 1020  
 DB 961 GCTATTACAGCAGAGAACGAGTATTAATACATCAAAAGACAGATGAATGAGAA 1020  
 OY 1021 AGGCTGAATTCGTAAATTAAGCAAAAGAGCTGTGAGCAAGAGCCCAATTAACAGAT 1080  
 DB 1021 AGGCTGAATTCGTAAATTAAGCAAAAGAGCTGTGAGCAAGAGCCCAATTAACAGAT 1080  
 OY 1081 GGGCTGCAAGTAAGAAACATGTATGATAGCGGACTCCAGACACAGAAAAAAGTAG 1140  
 DB 1081 GGGCTGCAAGTAAGAAACATGTATGATAGCGGACTCCAGACACAGAAAAAAGTAG 1140  
 OY 1141 ATTCGAATCTGATCCCTGCTGTGAGAGAAAGATGAATTAAGCAAACTGCATGCT 1200  
 DB 1141 ATTCGAATCTGATCCCTGCTGTGAGAGAAAGATGAATTAAGCAAACTGCATGCT 1200  
 OY 1201 CAGAGAAATCTAGAGATTAAGATGTTCTTGTGATTAACCTAATAGCAATTCAGA 1260  
 DB 1201 CAGAGAAATCTAGAGATTAAGATGTTCTTGTGATTAACCTAATAGCAATTCAGA 1260  
 OY 1261 AAGTTAATGAGTGTTCAGAGAGTGAATGATAGTTAGTTCTGATGACCTCACATGATG 1320  
 DB 1261 AAGTTAATGAGTGTTCAGAGAGTGAATGATAGTTAGTTCTGATGACCTCACATGATG 1320  
 OY 1321 GGGAGTCTGATCAAAAGCCAAAGTACCTGATGATGAGAGTGTCTAAATCAGGTAGATG 1380  
 DB 1321 GGGAGTCTGATCAAAAGCCAAAGTACCTGATGATGAGAGTGTCTAAATCAGGTAGATG 1380  
 OY 1381 AATATCTGCTTCTCAGAGAAATAGACTTACGACAGATGATCTCATAGAGCTTTAA 1440

```
1381 AATATCTGTTCTTCAGAGAAATAGACTTACGCGACGATCCTCATGAGCTTAA 1440
1441 TATGTAAAGTGAAGGTCACCCCAATCAGTAGAGATAATATGAAGCAAAATAT 1500
1441 TATGTAAAGTGAAGGTCACCCCAATCAGTAGAGATAATATGAAGCAAAATAT 1500
1501 TTGGGAAAACCTTTCGGAAGAGCAAGCCCTCCCACTTAAGCCATGTAAGTGAATC 1560
1501 TTGGGAAAACCTTTCGGAAGAGCAAGCCCTCCCACTTAAGCCATGTAAGTGAATC 1560
1501 TTGGGAAAACCTTTCGGAAGAGCAAGCCCTCCCACTTAAGCCATGTAAGTGAATC 1560
1561 TAAATTATGAGCATTTGTTACTGAGCCACAGATTAATACAGAGCGTCCCTCAANA 1620
1561 TAAATTATGAGCATTTGTTACTGAGCCACAGATTAATACAGAGCGTCCCTCAANA 1620
1621 AATTAAGGCTAAAGAGACCTACATCAGGCTTCCTCCTGAGATTTTATCAAGAA 1680
1621 AATTAAGGCTAAAGAGACCTACATCAGGCTTCCTCCTGAGATTTTATCAAGAA 1680
1681 CAGATTGGCAGTTCAAAAAGACTCTGGAATGATTAATCAGGAACTTACCAAGGAG 1740
1681 CAGATTGGCAGTTCAAAAAGACTCTGGAATGATTAATCAGGAACTTACCAAGGAG 1740
1741 AGAATGCTCAGTATGATATTTACTATATAGTGCTCATGAGATTAACAAAGGAT 1800
1741 AGAATGCTCAGTATGATATTTACTATATAGTGCTCATGAGATTAACAAAGGAT 1800
1801 CTATTCAGATGAGAAATCTTACCCATAGAAATAGTACGAAAGAAATCTGCTTCA 1860
1801 CTATTCAGATGAGAAATCTTACCCATAGAAATAGTACGAAAGAAATCTGCTTCA 1860
1861 AAACGAAGCTGAGCTATATAGAGCAGATTAAGCAATATGAACTGAAATTAATATCC 1920
1861 AAACGAAGCTGAGCTATATAGAGCAGATTAAGCAATATGAACTGAAATTAATATCC 1920
1921 ACAATTCAAAAGCACTTAAAGAAATAGGCTGAGAGGAAAGTCTTACCAAGCATATTC 1980
1921 ACAATTCAAAAGCACTTAAAGAAATAGGCTGAGAGGAAAGTCTTACCAAGCATATTC 1980
1981 ATGCGCTGTAAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGTAT 2040
1981 ATGCGCTGTAAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGTAT 2040
2041 TTGATAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGTAT 2100
2041 TTGATAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGTAT 2100
2101 GGCACAGCAGAAACCTTCAACATCATGAGAGTAAAGAACTGCAACTGGAGCCAAAG 2160
2101 GGCACAGCAGAAACCTTCAACATCATGAGAGTAAAGAACTGCAACTGGAGCCAAAG 2160
2161 GTTAAAGCCCAATGAGACAGCAAGTAAAGACATGAGTACTTCCCGAGAGTGA 2220
2161 GTTAAAGCCCAATGAGACAGCAAGTAAAGACATGAGTACTTCCCGAGAGTGA 2220
2221 AGTTAAACAATGCACTGCTTCTTACTAAGTGTTCCTCAATACAGTGAATTAAGAA 2280
2221 AGTTAAACAATGCACTGCTTCTTACTAAGTGTTCCTCAATACAGTGAATTAAGAA 2280
2281 TTGTCAATCTGACCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2340
2281 TTGTCAATCTGACCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2340
2281 TTGTCAATCTGACCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2340
2281 TTGTCAATCTGACCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2340
2341 GTTAAATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2400
2341 GTTAAATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2400
2401 AAGATCTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2460
2401 AAGATCTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2460
2461 AAGATCTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520
2461 AAGATCTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520
```

---

```
D 2461 AAGATCTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520
Q 2521 GTGTGAGTCAAGTGGAGAGATTTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580
D 2521 GTGTGAGTCAAGTGGAGAGATTTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580
Q 2581 ATATATAGAAATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2640
D 2581 ATATATAGAAATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2640
Q 2641 GGGAAACACATAGAAATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2700
D 2641 GGGAAACACATAGAAATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2700
Q 2701 TCAAGGTTTCAAGGCGCAGTATTTGCTGCTGCTTCAATCAGAGAGAGAGAGAGAG 2760
D 2701 TCAAGGTTTCAAGGCGCAGTATTTGCTGCTGCTTCAATCAGAGAGAGAGAGAGAG 2760
Q 2761 AATGTCAGATCTCTGCGCAGTCTGCTGCTTAAAGAGAGAGAGAGAGAGAGAGAGAG 2820
D 2761 AATGTCAGATCTCTGCGCAGTCTGCTGCTTAAAGAGAGAGAGAGAGAGAGAGAGAG 2820
Q 2821 TTGATATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2880
D 2821 TTGATATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2880
Q 2881 AGACAGTATATACAGTGGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2940
D 2881 AGACAGTATATACAGTGGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2940
Q 2941 ATGCAAAATGATATCAAGAGAGGCTGAGTTTGTCTATCATCTCAGTTCAGAGCA 3000
D 2941 ATGCAAAATGATATCAAGAGAGGCTGAGTTTGTCTATCATCTCAGTTCAGAGCA 3000
Q 3001 ACGAATGAGCATTACTCCCAATATACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060
D 3001 ACGAATGAGCATTACTCCCAATATACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060
Q 3061 CACCACCTTTTCCCATCAAGTCAATTTGTTAAACCTTAAATGAGAGAGAGAGAGAG 3120
D 3061 CACCACCTTTTCCCATCAAGTCAATTTGTTAAACCTTAAATGAGAGAGAGAGAGAG 3120
Q 3121 AAAACCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3180
D 3121 AAAACCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3180
Q 3181 GTACAGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3240
D 3181 GTACAGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3240
Q 3241 CAAGCAATATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3300
D 3241 CAAGCAATATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3300
Q 3301 TAGGTTCCAGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3360
D 3301 TAGGTTCCAGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3360
Q 3361 ATGCTATGCTTATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3420
D 3361 ATGCTATGCTTATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3420
Q 3421 GTTAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3480
D 3421 GTTAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3480
Q 3481 ATACAGATCTCTGCAATATCTGATTAATCAAGAGAGAGAGAGAGAGAGAGAGAGAG 3540
D 3481 ATACAGATCTCTGCAATATCTGATTAATCAAGAGAGAGAGAGAGAGAGAGAGAGAG 3540
Q 3541 ATGATCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3600
D 3541 ATGATCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3600
```

OY 3601 AACATCTAGTTTGGTGAAGAAATGACATTAGGAAGTTCTGCTGTTTTAGCAAAAGCG 3660  
|||||  
Db 3601 AAGATCTAGTTTGGTGAAGAAATGACATTAGGAAGTTCTGCTGTTTTAGCAAAAGCG 3660  
OY 3661 TCCGAAAGAGAGCTTAGCAGAGTCTAGCCCTTTCACCCATACATATTTGGCTCAG 3720  
|||||  
Db 3661 TCCGAAAGAGAGCTTAGCAGAGTCTAGCCCTTTCACCCATACATATTTGGCTCAG 3720  
OY 3721 GTTACGGAAGAGGGGCCAAGAAATTAGATCTCCAGAAAGAACTTATCTAGAGAGT 3780  
|||||  
Db 3721 GTTACGGAAGAGGGGCCAAGAAATTAGATCTCCAGAAAGAACTTATCTAGAGAGT 3780  
OY 3781 AAGAGCTTCCCTGCTCCAAACCTTGTATTGGTGAAGTAAATATACCTTCCAGT 3840  
|||||  
Db 3781 AAGAGCTTCCCTGCTCCAAACCTTGTATTGGTGAAGTAAATATACCTTCCAGT 3840  
OY 3841 CTACTAGGCAATAGCACCGTTGCTACCGAGTGTCTGTAGAAACACAGAGAGAAATTTAT 3900  
|||||  
Db 3841 CTACTAGGCAATAGCACCGTTGCTACCGAGTGTCTGTAGAAACACAGAGAGAAATTTAT 3900  
OY 3901 TATCATTTGAAGAAATACCTTAATGATGAGTAAACAGAGTAAATTTGGCAAGGCTCTC 3960  
|||||  
Db 3901 TATCATTTGAAGAAATACCTTAATGATGAGTAAACAGAGTAAATTTGGCAAGGCTCTC 3960  
OY 3961 AGGAACATCACCTTAGTGAAGAAACAAATGTTCTGCTAGCTGTTTCTCAGCTGGA 4020  
|||||  
Db 3961 AGGAACATCACCTTAGTGAAGAAACAAATGTTCTGCTAGCTGTTTCTCAGCTGGA 4020  
OY 4021 GTGAATTTGAAGACTTGACTGCAAAATACAAACACCCAGATCCTTCTGATGTTGTTCT 4080  
|||||  
Db 4021 GTGAATTTGAAGACTTGACTGCAAAATACAAACACCCAGATCCTTCTGATGTTGTTCT 4080  
OY 4081 CCAAAACAAATGAGGCACTCACTGGAAGCCAGGAGTTGCTGAGTGCACAGGAATTTGG 4140  
|||||  
Db 4081 CCAAAACAAATGAGGCACTCACTGGAAGCCAGGAGTTGCTGAGTGCACAGGAATTTGG 4140  
OY 4141 TTTCAGATGATGAAGAAAGAGAAAGCGGCTTGAAGAAATATATCAAGAAAGCAAGCA 4200  
|||||  
Db 4141 TTTCAGATGATGAAGAAAGAGAAAGCGGCTTGAAGAAATATATCAAGAAAGCAAGCA 4200  
OY 4201 TGGATTCAAACTTAGTGAAGCAGCATCTGGGTGTAGAGTGAAGCAAGCGTCTCGAAG 4260  
|||||  
Db 4201 TGGATTCAAACTTAGTGAAGCAGCATCTGGGTGTAGAGTGAAGCAAGCGTCTCGAAG 4260  
OY 4261 ACTGCTCAGGGCTATCTCTCAGAGTGAATTTTAACTCAGCAGAGAGGATACCATGC 4320  
|||||  
Db 4261 ACTGCTCAGGGCTATCTCTCAGAGTGAATTTTAACTCAGCAGAGAGGATACCATGC 4320  
OY 4321 AACATTAACCTGATTAAGCTCCAGAGGAATGGCTGAAGTGAAGCTGTGTTGAACAGC 4380  
|||||  
Db 4321 AACATTAACCTGATTAAGCTCCAGAGGAATGGCTGAAGTGAAGCTGTGTTGAACAGC 4380  
OY 4381 ATGGAGCAGAGCTTCTTAACAGCTACCTTCCATTAAGTGACTCTTCTGCCCTTGAG 4440  
|||||  
Db 4381 ATGGAGCAGAGCTTCTTAACAGCTACCTTCCATTAAGTGACTCTTCTGCCCTTGAG 4440  
OY 4441 ACCTGGGAATCCAGAAACAAAGCACATCGAAAAACAGATTAATTAACCTTCACAGAAAGTA 4500  
|||||  
Db 4441 ACCTGGGAATCCAGAAACAAAGCACATCGAAAAACAGATTAATTAACCTTCACAGAAAGTA 4500  
OY 4501 GTGAATTAACCTTAATTAAGCTCCAGAGGCTTCTGCTGCAAGTTTGAAGTCTG 4560  
|||||  
Db 4501 GTGAATTAACCTTAATTAAGCTCCAGAGGCTTCTGCTGCAAGTTTGAAGTCTG 4560  
OY 4561 CAGATAGTTCTACCACTAAATAAAGAACACAGAGTGAAGAGTATCCCTTCTTAAT 4620  
|||||  
Db 4561 CAGATAGTTCTACCACTAAATAAAGAACACAGAGTGAAGAGTATCCCTTCTTAAT 4620  
OY 4621 GCCCATCATTAAGTATAGTGTGTACATGACAGTGTGCTGTGGAGTCTTTCAGAAATGA 4680  
|||||  
Db 4621 GCCCATCATTAAGTATAGTGTGTACATGACAGTGTGCTGTGGAGTCTTTCAGAAATGA 4680

OY 4681 ACTACCATCTCAAGAGAGCTCATTAAGTTTGTGATGTGAGAGCAACACTGGAG 4740  
|||||  
Db 4681 ACTACCATCTCAAGAGAGCTCATTAAGTTTGTGATGTGAGAGCAACACTGGAG 4740  
OY 4741 AGTCTGGGCCACACGATTTTGACGGAAACATCTTACTTGGCCAAAGCAGATCTAGAGGAA 4800  
|||||  
Db 4741 AGTCTGGGCCACACGATTTTGACGGAAACATCTTACTTGGCCAAAGCAGATCTAGAGGAA 4800  
OY 4801 CCCCTTACCTGGAATCTGGAATCAGCTCTCTCTGTATGACCCCTGAAATCTGATCTTCTG 4860  
|||||  
Db 4801 CCCCTTACCTGGAATCTGGAATCAGCTCTCTCTGTATGACCCCTGAAATCTGATCTTCTG 4860  
OY 4861 AAGACAGAGCCCGAGATGCTGCTGTTGGCAACATATACATCTTAAACCTCTGCAATGA 4920  
|||||  
Db 4861 AAGACAGAGCCCGAGATGCTGCTGTTGGCAACATATACATCTTAAACCTCTGCAATGA 4920  
OY 4921 AAGTCCCAATTTGAAGTTGCAAAATCTGCCAGAGTCCAGCTGCTCATACTG 4980  
|||||  
Db 4921 AAGTCCCAATTTGAAGTTGCAAAATCTGCCAGAGTCCAGCTGCTCATACTG 4980  
OY 4981 ATACTGCTGGGTATTAATGCAATGGAAGAGTGTGAGAGAGGAGAAAGCCAAATTTGAC 5040  
|||||  
Db 4981 ATACTGCTGGGTATTAATGCAATGGAAGAGTGTGAGAGAGGAGAAAGCCAAATTTGAC 5040  
OY 5041 CTTCACAGAAAGGCTCAACAAAGATGTCATGGTGTGCTGGCTGACCCCAAG 5100  
|||||  
Db 5041 CTTCACAGAAAGGCTCAACAAAGATGTCATGGTGTGCTGGCTGACCCCAAG 5100  
OY 5101 AATTTATGCTGTGTACAAAGTTTGCAGAAACACACATCACTTAACTAATCTAATTA 5160  
|||||  
Db 5101 AATTTATGCTGTGTACAAAGTTTGCAGAAACACACATCACTTAACTAATCTAATTA 5160  
OY 5161 CTGAAGAGCTACTCATGTTGTTATGAAGAAACAGATGCTGATGTTGTGTGAAGCGAC 5220  
|||||  
Db 5161 CTGAAGAGCTACTCATGTTGTTATGAAGAAACAGATGCTGATGTTGTGTGAAGCGAC 5220  
OY 5221 TGAATATTTTCTAGGAATTTGCGGAGGAAATGGGTAGTGAATTTCTGGGTACCC 5280  
|||||  
Db 5221 TGAATATTTTCTAGGAATTTGCGGAGGAAATGGGTAGTGAATTTCTGGGTACCC 5280  
OY 5281 AGTCTATTAAGAAAGAAATATCTGAATGAGCATATTTGAAGTCAAGAGAGATGTGG 5340  
|||||  
Db 5281 AGTCTATTAAGAAAGAAATATCTGAATGAGCATATTTGAAGTCAAGAGAGATGTGG 5340  
OY 5341 TCAATGGAAGAAACCCAAAGTCCAAAGGAGCAGAGCAAGAAATCCCAAGAGAAATATC 5400  
|||||  
Db 5341 TCAATGGAAGAAACCCAAAGTCCAAAGGAGCAGAGCAAGAAATCCCAAGAGAAATATC 5400  
OY 5401 TCAGGGGGCTAGAAATCTGTTGCTATGAGGCTTTCACCAACATGCCCCAGATCAACTGG 5460  
|||||  
Db 5401 TCAGGGGGCTAGAAATCTGTTGCTATGAGGCTTTCACCAACATGCCCCAGATCAACTGG 5460  
OY 5461 AATGATGTGACAGCTGTGTGTCTTCTGTGTGAAGAGCTTTCATCATTCACCTTGG 5520  
|||||  
Db 5461 AATGATGTGACAGCTGTGTGTCTTCTGTGTGAAGAGCTTTCATCATTCACCTTGG 5520  
OY 5521 GCACAGGTCCACCCCAATTTGTGTGTGTGAGCCCAATGCTGTGAGAGAGAAATGGCT 5580  
|||||  
Db 5521 GCACAGGTCCACCCCAATTTGTGTGTGTGAGCCCAATGCTGTGAGAGAGAAATGGCT 5580  
OY 5581 TCCATGCAATTTGGGAGATGTGTGAGGACCTGTGTGACCCGAGAGTGGGTGGGACA 5640  
|||||  
Db 5581 TCCATGCAATTTGGGAGATGTGTGAGGACCTGTGTGACCCGAGAGTGGGTGGGACA 5640  
OY 5641 GTGTAGCACTTACAGTGCAGAGCTTGACACCTTACCTGATATACCCAGATCCCCACA 5700  
|||||  
Db 5641 GTGTAGCACTTACAGTGCAGAGCTTGACACCTTACCTGATATACCCAGATCCCCACA 5700  
OY 5701 GCCACTACTGA 5711  
|||||  
Db 5701 GCCACTACTGA 5711

RESULT 12
AAT84840
ID AAT84840 standard; CDNA to mRNA; 5712 BP.
XX AAT84840;
XX 21-FEB-1998 (first entry)
DT Human breast and ovarian cancer susceptibility gene BRCA1.
XX
KW BRCA1; breast cancer; ovarian cancer; human;
XX tumour suppressor gene; gene therapy; receptor; ds.
XX Homo sapiens.
OS
XX
FH Key
FT Location/Qualifiers
FT 120..5711
FT CDS
XX /tag= a
FT /transl_except= (pos:1482..1484, aa:asp)
FT /transl_except= (pos:1581..1583, aa:ser)
FT /transl_except= (pos:1699..1901, aa:glu)
FT /transl_except= (pos:1902..1904, aa:leu)
FT /transl_except= (pos:1905..1907, aa:leu)
FT /transl_except= (pos:1908..1910, aa:leu)
FT /transl_except= (pos:1911..1913, aa:asn)
FT /transl_except= (pos:1914..1916, aa:ile)
FT /transl_except= (pos:1917..1919, aa:ile)
FT /transl_except= (pos:4476..4478, aa:val)
FT /transl_except= (pos:4479..4481, aa:val)
FT /transl_except= (pos:4482..4484, aa:glu)
FT /transl_except= (pos:4689..4691, aa:pro)
XX W09730108-AI.
XX
PD 21-AUG-1997.
XX
PF 19-FEB-1997; 97WO-US03340.
XX
PX 20-FEB-1996; 96US-0603753.
XX
PA (OVVA-) UNIV VANDERBILT.
PA (UNIV) UNIV WASHINGTON.
XX
PI Claire-King M, Holt JT, Jensen RA, Jettion TL, Page DL;
DR Robinson-Bentley CL, Szabo CI, Thompson ME;
DR P-PSDB; AAM23286.
XX
BRCA1 and BRCA2 tumour suppressor gene products - useful to inhibit
PT breast and ovarian cancer cell growth and tumorigenesis, or treat
PT gene linked hereditary or sporadic ovarian or breast cancer
XX
PS Claim 13; Page 54-63; 148bp; English.
XX
CC This sequence comprises a full-length BRCA1 cDNA. Genetic analysis
CC of familial and ovarian cancer indicates that BRCA1 is a tumour
CC suppressor gene. It encodes a 190 kDa protein (see AAM23286) that is
CC an inhibitor of the growth and proliferation of human breast and
CC ovarian cancer cells. DNA encoding the BRCA1 protein can thus be
CC used in gene therapy methods for the treatment of breast and
CC ovarian cancers. A purified BRCA1 protein can also be used to treat
CC these cancers and, since it is secreted, can be used to identify
CC the BRCA1 receptor and hence to identify BRCA1 protein-mimetic
CC agents which act on the receptor for use in breast and ovarian
CC cancer treatment. The BRCA2 gene (see AAT84841) and BRCA2 protein
CC (see AAM23287) have also been characterised. Methods are claimed for
CC the isolation of BRCA1 or BRCA2 receptors, and for treating or
CC preventing breast and (gene-linked hereditary or sporadic) ovarian
CC cancers using BRCA1 and BRCA2 proteins and genes.
XX
Sequence 5712 BP; 1956 A; 1099 G; 1274 G; 1383 T; 0 other;

Query Match	99.9% ; Score 5707.8 ; DB 18 ; Length 5712 ;
Best Local Similarity	100.0% ; Pred. No. 0 ;
Matches 5709 ;	Conservative 0 ; Mismatches 2 ; Indels 0 ; Gaps 0 ;
OY	1 AGCTCGCGVGAAGCTTCCTGGACCCCGCAGCAAGGCTGTGGGGTTCCTAGATTAACCTGGCC 60
Db	1 AGCTCGCTGAGATTCCTCGAGCCCGCCAGCCAGGCTGTGGGGTTCCTAGATTAACCTGGCC 60
OY	61 CCGGCCCTGAGAGGCGCTTACACCTCTGCTGGGTAAGTTCATTTGGAAACGAAAGAA 120
Db	61 CCGGCCCTGAGAGGCGCTTACACCTCTGCTGGGTAAGTTCATTTGGAAACGAAAGAA 120
OY	121 TGGATTATATCTGCTCTCGCTTGAAGAAAGTCAAAATGCTATTAATGCTATGCAAGAA 180
Db	121 TGGATTATATCTGCTCTGCGCTTGAAGAAAGTCAAAATGCTATTAATGCTATGCAAGAA 180
OY	181 TCTTTAAGAGTGCCTCATCTGTCTGGAGTTGATCAAGAAACCTGTCTCCACAAGGTGACC 240
Db	181 TCTTTAAGAGTGCCTCATCTGTGCGCTTGAAGAAAGTCAAAATGCTATTAATGCTATGCAAGAA 240
OY	241 ACATATTTTGGCAAAATTTTGCATGCGAAGACTTTCACACCGAAGAAAGGCGCTTCACGT 300
Db	241 ACATATTTTGGCAAAATTTTGCATGCGAAGACTTTCACACCGAAGAAAGGCGCTTCACGT 300
OY	301 GTCCTTTATGTGAAGATGATATACCAAGAGGCTTACAGAAAGTACGAGATTTAGTC 360
Db	301 GTCCTTTATGTGAAGATGATATACCAAGAGGCTTACAGAAAGTACGAGATTTAGTC 360
OY	361 AACCTGTGTAGAGCATTTGAAGAAATCATTTGGTTTTCAGCTTTCACAGAGGTTGGAGT 420
Db	361 AACCTGTGTAGAGCATTTGAAGAAATCATTTGGTTTTCAGCTTTCACAGAGGTTGGAGT 420
OY	421 ATGCAAAACAGCTATATAATTTTGGCAAAAAGAAATAACTCTCCTGAAACATCTTAAGAAATG 480
Db	421 ATGCAAAACAGCTATATAATTTTGGCAAAAAGAAATAACTCTCCTGAAACATCTTAAGAAATG 480
OY	481 AATTTTCTATCATCCAAAGTATGGGCTACAGAAACCGTGGCAAAAGACTTTCACAGAGT 540
Db	481 AATTTTCTATCATCCAAAGTATGGGCTACAGAAACCGTGGCAAAAGACTTTCACAGAGT 540
OY	541 AACCCGAAAAATCCTTCTCTGAGAAACCAAGTCACTGTCCAACTCTAACCTTTGGAA 600
Db	541 AACCCGAAAAATCCTTCTCTGAGAAACCAAGTCACTGTCCAACTCTAACCTTTGGAA 600
OY	601 CTGTGGAACCTGTGAGACAAAGACGGGATACACTCAAAAGAGCTGTCTACATTTG 660
Db	601 CTGTGGAACCTGTGAGACAAAGACGGGATACACTCAAAAGAGCTGTCTACATTTG 660
OY	661 AATTTGGATCTGATTTCTTCTGGAAGTACCGTTAATAAGCAACTTTTTCAGTGGGAG 720
Db	661 AATTTGGATCTGATTTCTTCTGAGATACCGTTAATAAGCAACTTTTTCAGTGGGAG 720
OY	721 ATCAGAATTTGTACAAATCAACCCCTCAAGGAACAGGAGTGAATCACTTTGATTTGTG 780
Db	721 ATCAGAATTTGTACAAATCAACCCCTCAAGGAACAGGAGTGAATCACTTTGATTTGTG 780
OY	841 CCAGTAAATATATTTTGAACCACTAGAGAAAGCGTGCAGTGCAGAGCAAGT 900
Db	841 CCAGTAAATATATTTTGAACCACTAGAGAAAGCGTGCAGTGCAGAGCAATCCAGAAAGT 900
OY	901 ATCAGGGTATGTTCTGTTTCAAACTTGCATGTGGAGCCATGTGGCACAAATTAATCTCATGCCA 960
Db	901 ATCAGGGTATGTTCTGTTTCAAACTTGCATGTGGAGCCATGTGGCACAAATTAATCTCATGCCA 960
OY	961 GCTCATATACAGATAGACAGCAAGTTTATTAATCTCAATTAAGCAACAATGAATGTAGAA 1020
Db	961 GCTCATATACAGATAGACAGCAAGTTTATTAATCTCAATTAAGCAACAATGAATGTAGAA 1020



QY	1021	AGCGTGATTCGTGAATTAAGCAAAACCGCTGGCTTAGCAAGGACCAATCAACGAT	1089
Db	1021	AGCGTGATTCGTGAATTAAGCAAAACCGCTGGCTTAGCAAGGACCAATCAACGAT	1089
QY	1081	GGCGTGGAGTAAGGAACAACGTATGATGAGCGACCTCCAGCAGCAAGAAAAAGGTAG	1140
Db	1081	GGCGTGGAGTAAGGAACAACGTATGATGAGCGACCTCCAGCAGCAAGAAAAAGGTAG	1140
QY	1141	ATCTGAATGCTGATCCCGCTGTGTGAGAGAAAAGAAATGAATTAAGCAGAAACTGGCATGCT	1200
Db	1141	ATCTGAATGCTGATCCCGCTGTGTGAGAGAAAAGAAATGAATTAAGCAGAAACTGGCATGCT	1200
QY	1201	CAGGAATTCCTAGAGATCTGASAAATGTCTCTTGATTAACACTTAATAGCAGCTTACGA	1260
Db	1201	CAGGAATTCCTAGAGATCTGASAAATGTCTCTTGATTAACACTTAATAGCAGCTTACGA	1260
QY	1261	AAGTTAATGATGGGTTTTCCAGAGTGATGAACTGTTAGCTTGATGACTCACATGATG	1320
Db	1261	AAGTTAATGATGGGTTTTCCAGAGTGATGAACTGTTAGCTTGATGACTCACATGATG	1320
QY	1321	GGGAGTCTGAATCAAAATGCCAAAGTAGCTGATGTATTGGAGCTCTTAATAGGTTAGATG	1380
Db	1321	GGGAGTCTGAATCAAAATGCCAAAGTAGCTGATGTATTGGAGCTCTTAATAGGTTAGATG	1380
QY	1381	AATATTCTGGTCTTCACAGAGAAATAGACTTACTGGCCAGTAGTCCTCATGAGGCTTTAA	1440
Db	1381	AATATTCTGGTCTTCACAGAGAAATAGACTTACTGGCCAGTAGTCCTCATGAGGCTTTAA	1440
QY	1441	TATGTAAAGGGAAGAGTTCACTCCCAATCAGTAGAGAGTAATTAAGAGACAAATAT	1500
Db	1441	TATGTAAAGGGAAGAGTTCACTCCCAATCAGTAGAGAGTAATTAAGAGACAAATAT	1500
QY	1501	TTGGGAAAACCTATCGAGAGAAAGGCAAGCCCTCCCACTTAAGCCATGTAACCTAAATATC	1560
Db	1501	TTGGGAAAACCTATCGAGAGAAAGGCAAGCCCTCCCACTTAAGCCATGTAACCTAAATATC	1560
QY	1561	TAATTATAGAGACCTTTGTCTTACTGAGCCACAGATTAATACAAGAGCGTCCCTACAATA	1620
Db	1561	TAATTATAGAGACCTTTGTCTTACTGAGCCACAGATTAATACAAGAGCGTCCCTACAATA	1620
QY	1621	AATTAAAGCCTAAAGAGAGACCTCATCAGAGCCCTCACTCGAGAGATTTTATACAAGAAAG	1680
Db	1621	AATTAAAGCCTAAAGAGAGACCTCATCAGAGCCCTCACTCGAGAGATTTTATACAAGAAAG	1680
QY	1681	CAGATTTGGCAGTTCAAAAGACCTCTGAATATGATTAATCAAGGAACTTAACCAAAAGGAGC	1740
Db	1681	CAGATTTGGCAGTTCAAAAGACCTCTGAATATGATTAATCAAGGAACTTAACCAAAAGGAGC	1740
QY	1741	AGAATGCTCAAGTCATTAATTAATTAATTAATGCTGCTATGAGATTAATAACAAAGGTGATT	1800
Db	1741	AGAATGCTCAAGTCATTAATTAATTAATTAATTAATGCTGCTATGAGATTAATAACAAAGGTGATT	1800
QY	1801	CTATTTCGAATGAGAAAATCCTTAACCCATATGATTAATCTCGAAAAAGAAATCTGCTTTCA	1860
Db	1801	CTATTTCGAATGAGAAAATCCTTAACCCATATGATTAATCTCGAAAAAGAAATCTGCTTTCA	1860
QY	1861	AAAAGGAAGCTGAACCTATTAAGCAGCAGTATTAAGCAATATGAACTGCAATTTAAATATCC	1920
Db	1861	AAAAGGAAGCTGAACCTATTAAGCAGCAGTATTAAGCAATATGAACTGCAATTTAAATATCC	1920
QY	1921	ACAATTTCAAAAGCAGCTTAAGAAATGAGCTGAGAGGAAAGTCTTCTTACAGGACATATTC	1980
Db	1921	ACAATTTCAAAAGCAGCTTAAGAAATGAGCTGAGAGGAAAGTCTTCTTACAGGACATATTC	1980
QY	1981	ATGGCGTTGAACATGATGTGAGTAGAAATCTTAAGCCACCTAATTTGTATCGAATTTGCAAA	2040
Db	1981	ATGGCGTTGAACATGATGTGAGTAGAAATCTTAAGCCACCTAATTTGTATCGAATTTGCAAA	2040
QY	2041	TTGATTAATTTCTTACAGAGTGAAGAGATTAAGAAAAAAAGTACAAACCAATGCGACTCA	2100
Db	2041	TTGATTAATTTCTTACAGAGTGAAGAGATTAAGAAAAAAAGTACAAACCAATGCGACTCA	2100
QY	2101	GGCAGCAGCAAACTTAACAATCATGGAAGGTAAAGAACTCGCAACTGAGCCAAAGAAGA	2160

Db	2101	GGCAGCAGGAAACCTTACACTCATGGAAGGTAAGAACTCGCAACTGGAACCAAGAA	2160
QY	2161	GTAAACAAACCAAATGMAACAGACAGTAAGAACAATGACAGTACTCTCCAGAGCTGA	2220
Db	2161	GTAACAAAGCCAAATGACAGACAACTAAAGAATGACAGAGGATCTCTCCAGAGCTGA	2220
QY	2221	AGTTAACAAATGACACCTGCTCTTTTACTAACTGTTCAATATCCAGTAACTTAAAGAT	2280
Db	2221	AGTTAACAAATGACACCTGCTCTTTTACTAACTGTTCAATATCCAGTAACTTAAAGAT	2280
QY	2281	TTGTCATATCCTAGCCTTCCAGAGAGAAGAAAGAAAGAAACTAGAAACAGTTAAAGTCT	2340
Db	2281	TTGTCATATCCTAGCCTTCCAGAGAGAAGAAAGAAAGAAACTAGAAACAGTTAAAGTCT	2340
QY	2341	CTAATATATGCTGAAGAGACCCCAAGATCTCATGTTTAAGTGGAGAAAGGGTTTGGAACTG	2400
Db	2341	CTAATATATGCTGAAGAGACCCCAAGATCTCATGTTTAAGTGGAGAAAGGGTTTGGAACTG	2400
QY	2401	AAAGATCTGTAGAGAGTAGACAGTATTTTCATTGGTACCTGTTACTGATTATGGAAGCTAGG	2460
Db	2401	AAAGATCTGTAGAGAGTAGACAGTATTTTCATTGGTACCTGTTACTGATTATGGAAGCTAGG	2460
QY	2461	AAAGTATCTCGTTACTGGAAGTTAGACCTCTAGGAGAGCGCAAAAACAGAACCAATTAAT	2520
Db	2461	AAAGTATCTCGTTACTGGAAGTTAGACCTCTAGGAGAGCGCAAAAACAGAACCAATTAAT	2520
QY	2521	GTGTGAGTCACGTGTCACAGCATTTTGGAAACCCCAAGGGACTAATTCATGCTGTTCCAAAG	2580
Db	2521	GTGTGAGTCACGTGTCACAGCATTTTGGAAACCCCAAGGGACTAATTCATGCTGTTCCAAAG	2580
QY	2581	ATATAGAAATGACACAGAGGCTTTTAATGATTCACATTTGGGACATGAAGTTAAACACAGTC	2640
Db	2581	ATATAGAAATGACACAGAGGCTTTTAATGATTCACATTTGGGACATGAAGTTAAACACAGTC	2640
QY	2641	GGGAACCAAGCATAGAAATGGAAGAAGTAAGTAAGTCACTATTTGGCAGATACAT	2700
Db	2641	GGGAACCAAGCATAGAAATGGAAGAAGTAAGTAAGTCACTATTTGGCAGATACAT	2700
QY	2701	TCAGAGTTTCAAGAGCGCCAGTCATTGGTCTGTTTTCAATCCAGGAATCCAGAAAGG	2760
Db	2701	TCAGAGTTTCAAGAGCGCCAGTCATTGGTCTGTTTTCAATCCAGGAATCCAGAAAGG	2760
QY	2761	AATGTGCACACATTTCTTGCCCACTCTGSGGTCCTTAAAGAAACAAAGTCCAAAGTCACTT	2820
Db	2761	AATGTGCACACATTTCTTGCCCACTCTGSGGTCCTTAAAGAAACAAAGTCCAAAGTCACTT	2820
QY	2821	TTGATATGTGAACAAAGAGAAAGAAATCAAGAAAGCATGTATATATCAAGCCTGTAC	2880
Db	2821	TTGATATGTGAACAAAGAGAAAGAAATCAAGAAAGCATGTATATATCAAGCCTGTAC	2880
QY	2881	AGACAGTTAATATACATGACGAGGCTTCTGCTGGTGGTGGTGCACAAAGATTAAGCAGTTGATA	2940
Db	2881	AGACAGTTAATATACATGACGAGGCTTCTGCTGGTGGTGGTGCACAAAGATTAAGCAGTTGATA	2940
QY	2941	ATGCCAAATGTAGTATCAAGAGAGGCTAGGTTTGTCTATCAATCTCAGTTCCAGAGCA	3000
Db	2941	ATGCCAAATGTAGTATCAAGAGAGGCTAGGTTTGTCTATCAATCTCAGTTCCAGAGCA	3000
QY	3001	ACGAAACTGACACTATTAATCCAAATTAACATGAGACTTTTACAAACCCCATATCTGTATAC	3060
Db	3001	ACGAAACTGACACTATTAATCCAAATTAACATGAGACTTTTACAAACCCCATATCTGTATAC	3060
QY	3061	CACCACTTTTCCCATCAAGTCATTTTGTATAACTTAATGTATAGAAAAATGTCTAGAGG	3120
Db	3061	CACCACTTTTCCCATCAAGTCATTTTGTATAACTTAATGTATAGAAAAATGTCTAGAGG	3120
QY	3121	AAAACTTTGAGAGACATTCATATGTACCTGAAAGAAATGGAATGAGAAACATTCGCA	3180
Db	3121	AAAACTTTGAGAGACATTCATATGTACCTGAAAGAAATGGAATGAGAAACATTCGCA	3180
QY	3181	GTACAGTGACACATTAAGCCGTAATAACATTTAGAGAAATGTTTTTAAAGAACCCAGCT	3240
Db	3181	GTACAGTGACACATTAAGCCGTAATAACATTTAGAGAAATGTTTTTAAAGAACCCAGCT	3240



Db	3181	GTACAGTGAACCAATTACCGCTAATTAACATTAGAGAAAATGTTTTTAACAGACCACT	3240
Qy	3241	CAAGAAATATTATGAAGTAGGTTCCAGTACTAATGAAGTGGGCTCCAGTATTATGA	3300
Db	3241	CAAGCAATATTATGAAGTAGGTTCCAGTACTAATGAAGTGGGCTCCAGTATTATGA	3300
Qy	3301	TAGGTTCCAGTATATAAAACATTCAAGCAAGTATAGTAACACAGAGGGCCAAATTTGA	3360
Db	3301	TAGGTTCCAGTATATAAAACATTCAAGCAAGTATAGTAACACAGAGGGCCAAATTTGA	3360
Qy	3361	ATGCTATGCTTGATTAGGGGTTTTTGCAACCGTAGGCTCATTAACAAAGTCTCTGGAA	3420
Db	3361	ATGCTATGCTTGATTAGGGGTTTTTGCAACCGTAGGCTCATTAACAAAGTCTCTGGAA	3420
Qy	3421	GTAATTGTAAAGCATCCCTGCAAAATATAAAGCAAGATATGAAGTAGTTCAGACTGTA	3480
Db	3421	GTAATTGTAAAGCATCCCTGCAAAATATAAAGCAAGATATGAAGTAGTTCAGACTGTA	3480
Qy	3481	ATACAGATTTCTCTCCATATCTGATTAGATTAAGAACAGCTTATGGGAAGTAGTC	3540
Db	3481	ATACAGATTTCTCTCCATATCTGATTAGATTAAGAACAGCTTATGGGAAGTAGTC	3540
Qy	3541	ATGCAATCAGGTTTTGTTGCTGAGAACCTGATGACCTGTATGATGATGTTGAATTAAGG	3600
Db	3541	ATGCAATCAGGTTTTGTTGCTGAGAACCTGATGACCTGTATGATGATGTTGAATTAAGG	3600
Qy	3601	AAGATACTAGTTTTGCTGCAAAATGACATTAAGAAAGTTTGCCTGTTTTAGCAAAAGCG	3660
Db	3601	AAGATACTAGTTTTGCTGCAAAATGACATTAAGAAAGTTTGCCTGTTTTAGCAAAAGCG	3660
Qy	3661	TCCAGAAAGGAGACTTAGCAGAGAGTCTAGGCCCTTACCCATATACATTTGGCTCAGG	3720
Db	3661	TCCAGAAAGGAGACTTAGCAGAGAGTCTAGGCCCTTACCCATATACATTTGGCTCAGG	3720
Qy	3721	GTTACCGAAGAGGGGCCCAAAATTAAGTCTCTGAGAACACCTTATCTGTAGAGATG	3780
Db	3721	GTTACCGAAGAGGGGCCCAAAATTAAGTCTCTGAGAACACCTTATCTGTAGAGATG	3780
Qy	3781	AAGAGCTTCCCGCTTCCCAACACTGTTATTGTGTAAGTAACAAATATACCTTCTCAGT	3840
Db	3781	AAGAGCTTCCCGCTTCCCAACACTGTTATTGTGTAAGTAACAAATATACCTTCTCAGT	3840
Qy	3841	CTACTAGGCATAGCACCGTCTACCGAGTCTGTCTTAAGAACACAGAGAGAAATTTAT	3900
Db	3841	CTACTAGGCATAGCACCGTCTACCGAGTCTGTCTTAAGAACACAGAGAGAAATTTAT	3900
Qy	3901	TATCATTTAAAGTAATGCTTAATATGATCTGACATAACACAGTAAATATGGCAAGCATCTC	3960
Db	3901	TATCATTTAAAGTAATGCTTAATATGATCTGACATAACACAGTAAATATGGCAAGCATCTC	3960
Qy	3961	AGGAACATCACCTTAGTGAGGAAACAAATATTCTGTAGACTTGTCTTCTCACAGTGCA	4020
Db	3961	AGGAACATCACCTTAGTGAGGAAACAAATATTCTGTAGACTTGTCTTCTCACAGTGCA	4020
Qy	4021	GTGAATTCGAAGACTTGAAGTCAAAATACAAACACCCAGAGATCTTCTTGATTTGGTTCTT	4080
Db	4021	GTGAATTCGAAGACTTGAAGTCAAAATACAAACACCCAGAGATCTTCTTGATTTGGTTCTT	4080
Qy	4081	CCAAACAANTAGAGCATGATGTGAAACCCAGGAGTGGTCTGATGACAGAAATTTGG	4140
Db	4081	CCAAACAANTAGAGCATGATGTGAAACCCAGGAGTGGTCTGATGACAGAAATTTGG	4140
Qy	4141	TTTCAGATGATGAAAGAGAGAGGAGGCTTGGAAGAAATTAATCAAGAAAGCAAAAGCA	4200
Db	4141	TTTCAGATGATGAAAGAGAGAGGAGGCTTGGAAGAAATTAATCAAGAAAGCAAAAGCA	4200
Qy	4201	TGGATTTCAAACCTTAGGTGAAGCAGCATCTGGGGTGTAGAGTGAAGAAAGCGCTCTGAAG	4260
Db	4201	TGGATTTCAAACCTTAGGTGAAGCAGCATCTGGGGTGTAGAGTGAAGAAAGCGCTCTGAAG	4260
Qy	4261	ACTGCTCAGGGCTATCTCTTCAAGATGACATTTTAACCACTCAGCAGAGGATATCCATGC	4320
Db	4261	ACTGCTCAGGGCTATCTCTTCAAGATGACATTTTAACCACTCAGCAGAGGATATCCATGC	4320

QY	4321	AAACATPAACCTGATYAAAGCTCCAGCCAGGAATGCTGAACCTAGACGCTGTTAGAACAGC	4380
Db	4321	AAACATPAACCTGATYAAAGCTCCAGCCAGGAATGCTGAACCTAGAACGCTGTTAGAACAGC	4380
QY	4381	ATGGGAGCCAGCCTTCTTAACACGTCACCTCCATATAAGACATCTTCTGCTGGTGGAG	4440
Db	4381	ATGGGAGCCAGCCTTCTTAACACGTCACCTCCATATAAGACATCTTCTGCTGGTGGAG	4440
QY	4441	ACCTGCGAATATCCAGAACCAAGACATCAGAAAAGACAGTATTAACTTACAGAAAAGTA	4500
Db	4441	ACCTGCGAATATCCAGAACCAAGACATCAGAAAAGACAGTATTAACTTACAGAAAAGTA	4500
QY	4501	GTCGATTAACCTCTATAGCCAGATTCAGAAAGCCTTCTGCTGACAACTTTAGAGTCGTG	4560
Db	4501	GTCGATTAACCTCTATAGCCAGATTCAGAAAGCCTTCTGCTGACAACTTTAGAGTCGTG	4560
QY	4561	CAGATAGTCTTACAGATAAAAAATAAGAACCCAGAGTGGAAAGTCATCCCTCTTAAT	4620
Db	4561	CAGATAGTCTTACAGATAAAAAATAAGAACCCAGAGTGGAAAGTCATCCCTCTTAAT	4620
QY	4621	GCCCATCATTAATAGATGATAGGTGGTACATGACACAGTTGCTCTGGAGTCTTGAGATPAGAA	4680
Db	4621	GCCCATCATTAATAGATGATAGGTGGTACATGACACAGTTGCTCTGGAGTCTTGAGATPAGAA	4680
QY	4681	ACTTACCACATCTCAAGAGGAGCGCTCATTAAGTGTGTTGATGTGGAGAGCAACAGCTGGAG	4740
Db	4681	ACTTACCACATCTCAAGAGGAGCGCTCATTAAGTGTGTTGATGTGGAGAGCAACAGCTGGAG	4740
QY	4741	AGTCGTGGGCGACAGATTTTGACGGAACATCTTACTTCCCAAGGCAAGATCTAGAGGAA	4800
Db	4741	AGTCGTGGGCGACAGATTTTGACGGAACATCTTACTTCCCAAGGCAAGATCTAGAGGAA	4800
QY	4801	CCCTTACTTGAATCTGGAATCCAGCCCTCTCTGATAGACCCGTAATCTATCCTCTGTG	4860
Db	4801	CCCTTACTTGAATCTGGAATCCAGCCCTCTCTGATAGACCCGTAATCTATCCTCTGTG	4860
QY	4861	AAGACAGAGCCCCACAGTCAGCTCCTGTTGGCACATATCCATCTTCAACCTCTGCATTTGA	4920
Db	4861	AAGACAGAGCCCCACAGTCAGCTCCTGTTGGCACATATCCATCTTCAACCTCTGCATTTGA	4920
QY	4921	AAGTTTCCCCAATTGAAAGTTGCCAGATCTGCCACAGTCAGCTGCTGCTATACTG	4980
Db	4921	AAGTTTCCCCAATTGAAAGTTGCCAGATCTGCCACAGTCAGCTGCTGCTATACTG	4980
QY	4981	ATACTGCTGGGTATATATGCAATGGAAGAAAGTGTGAGCGAGGAGAGCCAGATTTGACAG	5040
Db	4981	ATACTGCTGGGTATATATGCAATGGAAGAAAGTGTGAGCGAGGAGAGCCAGATTTGACAG	5040
QY	5041	CTTCAACAGAAAGGTCACAACAAGAATGTCATGTGCTGTGGCCTGACCCCAAG	5100
Db	5041	CTTCAACAGAAAGGTCACAACAAGAATGTCATGTGCTGTGGCCTGACCCCAAG	5100
QY	5101	AATTATGCTGCTGATACAACTTTGGCACAAAAACACACATCATTAACTATATCTAATTA	5160
Db	5101	AATTATGCTGCTGATACAACTTTGGCACAAAAACACACATCATTAACTATATCTAATTA	5160
QY	5161	CTGAAGAGACTACTCATGTTGTTATGAAGACAGATGCTGAGTTTGTGTGAACGACAC	5220
Db	5161	CTGAAGAGACTACTCATGTTGTTATGAAGACAGATGCTGAGTTTGTGTGAACGACAC	5220
QY	5221	TGAATTAATTTCTTAGAATTTGGGAGGAAAAATGGGTAGTTAGCTATTCTTGGGTGACCC	5280
Db	5221	TGAATTAATTTCTTAGAATTTGGGAGGAAAAATGGGTAGTTAGCTATTCTTGGGTGACCC	5280
QY	5281	AGTCATTTAAAGAAAGAAAAATGCTGATATGAGCATATTTTGAATCTCAGAGAGATGTGG	5340
Db	5281	AGTCATTTAAAGAAAGAAAAATGCTGATATGAGCATATTTTGAATCTCAGAGAGATGTGG	5340
QY	5341	TCAATGGAAGAACCAACAGGTCCTCAAAAGCGACAGAGAAATCCAGGACAGAAAGATCT	5400
Db	5341	TCAATGGAAGAACCAACAGGTCCTCAAAAGCGACAGAGAAATCCAGGACAGAAAGATCT	5400

QY 5401 TCAGGGGGCTAGAAATCTGTTGCTATGAGGCCCTTACCACACATGCCACAGATCAACTGG 5460  
|||||  
Db 5401 TCAGGGGGCTAGAAATCTGTTGCTATGAGGCCCTTACCACACATGCCACAGATCAACTGG 5460  
QY 5461 AATGATGATGACAGCTGTGTGCTTCTGTGGTGAAGAGACTTTCATTCACCTTG 5520  
|||||  
Db 5461 AATGATGATGACAGCTGTGTGCTTCTGTGGTGAAGAGACTTTCATTCACCTTG 5520  
QY 5521 GCACAGGTGTCACCCCAATTTGGTGTGTGACGCCAGATGGCTGACAGAGACATGGCT 5580  
|||||  
Db 5521 GCACAGGTGTCACCCCAATTTGGTGTGTGACGCCAGATGGCTGACAGAGACATGGCT 5580  
QY 5581 TCCATGCAATTTGGGACAGATGTGTGAGGACCTGTGTGACCCGAGAGTGGTGTGGACA 5640  
|||||  
Db 5581 TCCATGCAATTTGGGACAGATGTGTGAGGACCTGTGTGACCCGAGAGTGGTGTGGACA 5640  
QY 5641 GTGAGACACTCTACCACTGCGCAGAGCTGCACACCTGATGATACCCAGATCCGCCACA 5700  
|||||  
Db 5641 GTGAGACACTCTACCACTGCGCAGAGCTGCACACCTGATGATACCCAGATCCGCCACA 5700  
QY 5701 GCCACTACTGA 5711  
|||||  
Db 5701 GCCACTACTGA 5711  
RESULT 13  
AA287995 standard; DNA; 5712 BP.  
ID AA287995  
AC AA287995;  
DT 06-JUN-2000 (first entry)  
DE BRCA1 gene sequence (Genbank Accn No: U14680).  
XX BRCA1: tumour; prostate cancer; cytostatic; antiproliferative;  
KW gene therapy; ds.  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 120..5711  
FT /\*lag= a  
FT /transl\_except= (pos: 1452..1454, aa:Asp)  
FT /transl\_except= (pos: 1482..1484, aa:Asp)  
FT /transl\_except= (pos: 1581..1583, aa:Ser)  
FT /transl\_except= (pos: 1899..1901, aa:Glu)  
FT /transl\_except= (pos: 1902..1904, aa:Leu)  
FT /transl\_except= (pos: 1905..1907, aa:Glu)  
FT /transl\_except= (pos: 1908..1910, aa:Leu)  
FT /transl\_except= (pos: 1911..1913, aa:Asp)  
FT /transl\_except= (pos: 1914..1916, aa:Leu)  
FT /transl\_except= (pos: 1917..1919, aa:Met)  
FT /transl\_except= (pos: 4476..4478, aa:Val)  
FT /transl\_except= (pos: 4479..4481, aa:Leu)  
FT /transl\_except= (pos: 4482..4484, aa:Glu)  
FT /transl\_except= (pos: 4533..4535, aa:Xaa)  
FT /transl\_except= (pos: 4689..4691, aa:Pro)  
FT /note= "Xaa= unknown"  
XX CA2218197-A1.  
XX 05-JUN-1999.  
XX 12-DEC-1997; 97CA-2218197.  
XX 05-DEC-1997; 97US-0986106.  
XX (UNIV ) UNIV WASHINGTON.  
XX (UYVA-) UNIV VANDERBILT.  
XX Robinson-Benion CL, Thompson ME, Holt JT, Jensen RA, Steiner MS;  
PI King M;

XX WPI: 2000-238071/21.  
DR P-PSDB: AA177818.  
XX  
PT New method of treatment and suppression of prostate cancer comprises  
PT using the BRCA family of genes to decrease the growth rate of the tumor  
PT  
PS Claim 11: Page 98-104; 166pp; English.  
XX  
XX The invention relates to a method for suppressing the growth of a  
CC prostate tumor in a mammal that comprises introducing to the tumor a  
CC vector comprising a nucleic acid sequence encoding a BRCA family gene  
CC product operatively linked to a promoter, where production of the BRCA  
CC family gene product results in a decrease in the growth rate of the  
CC tumor. The methods are used to suppress the growth of and also to treat  
CC prostate cancer in a mammal where the tumor is gene-linked hereditary  
CC prostate cancer or sporadic prostate cancer. The present sequence  
CC represents a BRCA1 gene sequence (GenBank Accn No: U14680) that can be  
CC administered for suppressing tumor.  
XX  
SQ Sequence 5712 BP; 1956 A; 1099 C; 1274 G; 1383 T; 0 other:  
Query Match 99.9%; Score 5707.8; DB 21; Length 5712;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 5709; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 AGCTGCGTGAGACTTCCCTGACACCCGACAGAGCTGTGGGTTCTCAGATTAAGTGGCC 60  
Db 1 AGCTGCGTGAGACTTCCCTGACACCCGACAGAGCTGTGGGTTCTCAGATTAAGTGGCC 60  
QY CCTGCGCTCAGAGAGGCGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120  
Db CCTGCGCTCAGAGAGGCGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120  
QY 121 TGGATTTATCTGCTCTTGGGCTTGAAGATACAAATGTCATTATGCTATGCAAGAAA 180  
Db 121 TGGATTTATCTGCTCTTGGGCTTGAAGATACAAATGTCATTATGCTATGCAAGAAA 180  
QY 181 TCTTAGAGTGTCCATCTGTCTGAGTGTGATCAAGAAACCTGTCTCCAAAGTGTACC 240  
Db 181 TCTTAGAGTGTCCATCTGTCTGAGTGTGATCAAGAAACCTGTCTCCAAAGTGTACC 240  
QY 241 ACATATTTTGAATTTTGAATGCTGAACCTTCCACCCGACCAAGAGGCGCTTCAAGT 300  
Db 241 ACATATTTTGAATTTTGAATGCTGAACCTTCCACCCGACCAAGAGGCGCTTCAAGT 300  
QY 301 GTCTTTATGTAAAGATGATATACCAAAAGAGAGCTTCAAGAAAGTACGATTTAGTC 360  
Db 301 GTCTTTATGTAAAGATGATATACCAAAAGAGAGCTTCAAGAAAGTACGATTTAGTC 360  
QY 361 AACTTGTGAAGAGCTTGAATTAATCAATTTGCTTTTCAAGCTTGAACAGGTTTGAAGT 420  
Db 361 AACTTGTGAAGAGCTTGAATTAATCAATTTGCTTTTCAAGCTTGAACAGGTTTGAAGT 420  
QY 421 ATGCAACAGCTATTAATTTTGAAGAAAGAAATTAAGTCTCCGACATCTTAAGATG 480  
Db 421 ATGCAACAGCTATTAATTTTGAAGAAAGAAATTAAGTCTCCGACATCTTAAGATG 480  
QY 481 AAGTTTCTATCAATCAAGATATGGCTTACAGAAACCGTGCAGAAAGCTTCTACAGAGTG 540  
Db 481 AAGTTTCTATCAATCAAGATATGGCTTACAGAAACCGTGCAGAAAGCTTCTACAGAGTG 540  
QY 541 AACCCGAAATCTCTCTTGCAGAAACAGCTCACTGTCACATCTTAACCTTGGAA 600  
Db 541 AACCCGAAATCTCTCTTGCAGAAACAGCTCACTGTCACATCTTAACCTTGGAA 600  
QY 601 CTGTGAGACTCTGAGAGCAAGAGGAGTACAACTTCAAGAGCTGTGTACATTTG 660  
Db 601 CTGTGAGACTCTGAGAGCAAGAGGAGTACAACTTCAAGAGCTGTGTACATTTG 660  
QY 661 AATTGGATCTGATTTCTTGAAGATACCGTTAATAGGCACTTATGCACTGTGGAG 720  
Db 661 AATTGGATCTGATTTCTTGAAGATACCGTTAATAGGCACTTATGCACTGTGGAG 720

Db 661 AATTGGGATCTGATCTCTTGGAAGATACGTTAATAAGCACTTATTGCACTGTGGAG 720  
QY 721 ATCAGAATTTGTACAATCACCCCTCAAGAACCCAGGATGAATCAGTTGGATCTG 780  
Db 721 ATCAGAATTTGTACAATCACCCCTCAAGAACCCAGGATGAATCAGTTGGATCTG 780  
QY 781 CAAAAAAGCTGCTTGTGAATTTTCTGAGACGGATGTAAACAATACTGAACATCATCAC 840  
Db 781 CAAAAAAGCTGCTTGTGAATTTTCTGAGACGGATGTAAACAATACTGAACATCATCAC 840  
QY 841 CCAGTATATATGTTTGAACACACCTGAGAGGGTGGAGTGAAGGATCCAGAAAGT 900  
Db 841 CCAGTATATATGTTTGAACACACCTGAGAGGGTGGAGTGAAGGATCCAGAAAGT 900  
QY 901 ATCAGGATAGTTCTGTTTCAAACTGTGATGTGAGGCAATGTGACAAATACTATGCA 960  
Db 901 ATCAGGATAGTTCTGTTTCAAACTGTGATGTGAGGCAATGTGACAAATACTATGCA 960  
QY 961 GCTCATTTACAGCATGGAACACAGATTTATTACTCACTAAAGACAGATGAATGTAGAA 1020  
Db 961 GCTCATTTACAGCATGGAACACAGATTTATTACTCACTAAAGACAGATGAATGTAGAA 1020  
QY 1021 AGGCTGAATTTCTGTATTAATAAGCAACAGCTGGCTTAGCAGAGGCCAATACAGAT 1080  
Db 1021 AGGCTGAATTTCTGTATTAATAAGCAACAGCTGGCTTAGCAGAGGCCAATACAGAT 1080  
QY 1081 GGGCTGGAAGTAAGAAACATGTATGATAGGCGGACCTCCAGCACAGAAAAAAGTAG 1140  
Db 1081 GGGCTGGAAGTAAGAAACATGTATGATAGGCGGACCTCCAGCACAGAAAAAAGTAG 1140  
QY 1141 ATCTGAATGCTGATCCCTGTGTGAGAGAAAGATGAATAGCAGAAACTGCCATGCT 1200  
Db 1141 ATCTGAATGCTGATCCCTGTGTGAGAGAAAGATGAATAGCAGAAACTGCCATGCT 1200  
QY 1201 CAGAGATCTGAGATGATGAGATGTTCTGATTAACATTAATAGAGATTCAGA 1260  
Db 1201 CAGAGATCTGAGATGATGAGATGTTCTGATTAACATTAATAGAGATTCAGA 1260  
QY 1261 AAGTATAGTGTGTTTCCAGAAAGTATGATGTTAGTGTGATGACTCAATGATG 1320  
Db 1261 AAGTATAGTGTGTTTCCAGAAAGTATGATGTTAGTGTGATGACTCAATGATG 1320  
QY 1321 GGGAGCTGATCAATGCAAAAGTATGATGTTGACGTTCTTAATGAGGTAGTG 1380  
Db 1321 GGGAGCTGATCAATGCAAAAGTATGATGTTGACGTTCTTAATGAGGTAGTG 1380  
QY 1381 AATATCTGCTTCTAGAGAAATAGACTTACTGCGACATGCTCATGAGGCTTTAA 1440  
Db 1381 AATATCTGCTTCTAGAGAAATAGACTTACTGCGACATGCTCATGAGGCTTTAA 1440  
QY 1441 TATGTAAAGTGAAGAGTTCACCTCAATCAGTAGAGATTAATTGAAGACAAATAT 1500  
Db 1441 TATGTAAAGTGAAGAGTTCACCTCAATCAGTAGAGATTAATTGAAGACAAATAT 1500  
QY 1501 TTGGGAAAACTATCGAAGAGGCAAGCTCCCAACTTAAGCCATGTAACGAAATC 1560  
Db 1501 TTGGGAAAACTATCGAAGAGGCAAGCTCCCAACTTAAGCCATGTAACGAAATC 1560  
QY 1561 TAAATATAGAGATTTGTTACTGAGCAGATATATACAAAGAGCGCCCTCACAATA 1620  
Db 1561 TAAATATAGAGATTTGTTACTGAGCAGATATATACAAAGAGCGCCCTCACAATA 1620  
QY 1621 AATTAAGCGTAAAGAGACCTACATCAGGCTTATCTGAGATTTTATCAAAAG 1680  
Db 1621 AATTAAGCGTAAAGAGACCTACATCAGGCTTATCTGAGATTTTATCAAAAG 1680  
QY 1681 CAGATTTGGCAGTTCAAAAGACTCTGAAATGATTAATCAGGACTTAACCAAGGAGC 1740  
Db 1681 CAGATTTGGCAGTTCAAAAGACTCTGAAATGATTAATCAGGACTTAACCAAGGAGC 1740  
QY 1741 AGAATGTCAGTGAATATTACTAATAGTGTATGATGAATTAACCAAAAGGATTT 1800  
Db 1741 AGAATGTCAGTGAATATTACTAATAGTGTATGATGAATTAACCAAAAGGATTT 1800  
QY 1801 CTATTCGAATGAGAAAAATCTTAACCAATAGAAATCACTCGAAAAAGATCTGTTCA 1860  
Db 1801 CTATTCGAATGAGAAAAATCTTAACCAATAGAAATCACTCGAAAAAGATCTGTTCA 1860  
QY 1861 AAACGAAGCTGACCTATATAGCAGATATAGCAATGTGAACTCGAATTAATATTC 1920  
Db 1861 AAACGAAGCTGACCTATATAGCAGATATAGCAATGTGAACTCGAATTAATATTC 1920  
QY 1921 ACAATTCAAAAGCACTTAATAAGATAGGCTGAGGAAGTCTTACACGATATTTC 1980  
Db 1921 ACAATTCAAAAGCACTTAATAAGATAGGCTGAGGAAGTCTTACACGATATTTC 1980  
QY 1981 ATGCGCTTGAACATGATAGTACAGTAAGATCTAAGCCCACTAATTGTACTGAATTC 2040  
Db 1981 ATGCGCTTGAACATGATAGTACAGTAAGATCTAAGCCCACTAATTGTACTGAATTC 2040  
QY 2041 TTGATAGTTTCTTACAGTGAAGACATTAAGAAAAAAGTACAAACCAATGCCATGCA 2100  
Db 2041 TTGATAGTTTCTTACAGTGAAGACATTAAGAAAAAAGTACAAACCAATGCCATGCA 2100  
QY 2101 GGCACAGCAGAACTTACAACTCATGAGGTAAAGACCTGCACTGAGCCAGAAAGA 2160  
Db 2101 GGCACAGCAGAACTTACAACTCATGAGGTAAAGACCTGCACTGAGCCAGAAAGA 2160  
QY 2161 GTAAACAGCCAAATGAAACAGACAGTAAAGACATGACATGATCTTCCAGAGCTGA 2220  
Db 2161 GTAAACAGCCAAATGAAACAGACAGTAAAGACATGACATGATCTTCCAGAGCTGA 2220  
QY 2221 AGTTAAACAATGACCTGCTCTTTTACTAAGTGTTCATTAATCCAGTCACTTAAGAT 2280  
Db 2221 AGTTAAACAATGACCTGCTCTTTTACTAAGTGTTCATTAATCCAGTCACTTAAGAT 2280  
QY 2281 TTGTCAATCTTACCTTCCAGAGAAAGAAAGAAAGAACTAGAAACAGTTAAAGT 2340  
Db 2281 TTGTCAATCTTACCTTCCAGAGAAAGAAAGAAAGAACTAGAAACAGTTAAAGT 2340  
QY 2341 CTAAATATGCTGAAGACCCCAAGATCTCATGTTAAGTGAAGAGGTTTTGCAACTG 2400  
Db 2341 CTAAATATGCTGAAGACCCCAAGATCTCATGTTAAGTGAAGAGGTTTTGCAACTG 2400  
QY 2401 AAAGATCTGAGAGATGACGATTTTCAATGCTGATCTGATCTGATTAATGACACTG 2460  
Db 2401 AAAGATCTGAGAGATGACGATTTTCAATGCTGATCTGATCTGATTAATGACACTG 2460  
QY 2461 AAAGATCTGAGAGATGACGATTTTCAATGCTGATCTGATCTGATTAATGACACTG 2520  
Db 2461 AAAGATCTGAGAGATGACGATTTTCAATGCTGATCTGATCTGATTAATGACACTG 2520  
QY 2521 GTGTAGTCAAGTGTGAGCATTTGAAAACCCCAAGGACTAATTCATGTTGTTCAAG 2580  
Db 2521 GTGTAGTCAAGTGTGAGCATTTGAAAACCCCAAGGACTAATTCATGTTGTTCAAG 2580  
QY 2581 ATTAATGAATGACACAGAAAGCTTAAGTATCCATTTGGGACATGAATTAACCAATC 2640  
Db 2581 ATTAATGAATGACACAGAAAGCTTAAGTATCCATTTGGGACATGAATTAACCAATC 2640  
QY 2641 GGGAAACAAGCATGAAGATGAAAGAAAGTGAATGATGCTAGATTTGGAGATACAT 2700  
Db 2641 GGGAAACAAGCATGAAGATGAAAGAAAGTGAATGATGCTAGATTTGGAGATACAT 2700  
QY 2701 TCAAGTTTCAAGGCCAGTCAATTTGCTGTTTCAATCCAGAAATGCAGAAAGG 2760  
Db 2701 TCAAGTTTCAAGGCCAGTCAATTTGCTGTTTCAATCCAGAAATGCAGAAAGG 2760  
QY 2761 AATGTCAACATTTCTGCCCACCTGCTGCTTAAGAACAAATGCCAAAGTCACTT 2820  
Db 2761 AATGTCAACATTTCTGCCCACCTGCTGCTTAAGAACAAATGCCAAAGTCACTT 2820  
QY 2821 TTGAATGTGAACAAAGAAAGAAATCAAGGAAAGAAAGTGAATGATCAAGCCTGTAC 2880  
Db 2821 TTGAATGTGAACAAAGAAAGAAATCAAGGAAAGAAAGTGAATGATCAAGCCTGTAC 2880

QY	2881	AGACAGTTAATATACACGACGAGCTTTCTCTGCGTGGTGGTGCACAAAGATATAGCCAGTTGATA	2940
Db	2881	AGACAGTTAATATACGACGAGCTTTCTCTGCGTGGTGGTGCACAAAGATATAGCCAGTTGATA	2940
QY	2941	ATGCCAAATGATGATCAAAAGGACGCTCTAGGTTTGTCTATCTCATCTCAGTTCAGAGCA	3000
Db	2941	ATGCCAAATGATGATCAAAAGGAGGCTCTAGGTTTGTCTATCTCATCTCAGTTCAGAGCA	3000
QY	3001	ACGAACCTGCACTTACTCTCCAAATAAAGATGACCTTTTACAAACCATTATCTGTATAC	3060
Db	3001	ACGAACCTGGACCTACTTACTCTCCAAATAAAGATGACCTTTTACAAACCATTATCTGTATAC	3060
QY	3061	CACCACTTTTCCCATCAAGTCATTTGTGTAAACTAAATGTAGAAAAATCTGCTAGAGG	3120
Db	3061	CACCACTTTTCCCATCAAGTCATTTGTGTAAACTAAATGTAGAAAAATCTGCTAGAGG	3120
QY	3121	AAACTTTGAGAACATTTCAATGTGCACCTGAGAAAGAAATGGGAAATGAGAACCTTCCAA	3180
Db	3121	AAACTTTGAGAACATTTCAATGTGCACCTGAGAAAGAAATGGGAAATGAGAACCTTCCAA	3180
QY	3181	GTACAGTGAGCACAAATTAGCCGTATATAACATTTAGAGAAAATGTTTTTAAAGAACCCAGCT	3240
Db	3181	GTACAGTGAGCACAAATTAGCCGTATATAACATTTAGAGAAAATGTTTTTAAAGAACCCAGCT	3240
QY	3241	CAACCACTATTATGAAGTAGTTCGCGTACTAATGAAGTGGGCTCCAGTATTATGAAA	3300
Db	3241	CAACCACTATTATTAAGAAAGTAGTTCGCGTACTAATGAAGTGGGCTCCAGTATTATGAAA	3300
QY	3301	TAGGTTCCAGTGAAGAAACCTTCACAGAACCTAGTAGAAACAGAGGGCCAAATTTGA	3360
Db	3301	TAGGTTCCAGTGAAGAAACCTTCACAGAACCTAGTAGAAACAGAGGGCCAAATTTGA	3360
QY	3361	ATGCTATGCTTAGATTAGGGGTTTTTGCACCTGAGGTCCTATAAACAAGTCTTCTCGAA	3420
Db	3361	ATGCTATGCTTAGATTAGGGGTTTTTGCACCTGAGGTCCTATAAACAAGTCTTCTCGAA	3420
QY	3421	GTAATTGTAAACATCTCGAAATTAAGAAACAGAAATATGAAGAGTAGTTCAGACTGTTA	3480
Db	3421	GTAATTGTAAACATCTCGAAATTAAGAAACAGAAATATGAAGAGTAGTTCAGACTGTTA	3480
QY	3481	ATACAGATTTCTCTCCATCTCTGATCTGATTAACCTTAGAACAGCCATATGGGAAGTAGTC	3540
Db	3481	ATACAGATTTCTCTCCATCTCTGATCTGATTAACCTTAGAACAGCCATATGGGAAGTAGTC	3540
QY	3541	ATGCATCTCAAGGTTGTTCTTGAGAGACACCTGATGACCTGTAGATGATGGTGAATTAAGG	3600
Db	3541	ATGCATCTCAAGGTTGTTCTTGAGAGACACCTGATGACCTGTAGATGATGGTGAATTAAGG	3600
QY	3601	AAGATCTAGTTTGGTGAANAATGACATTTAAGGAAAGTTCTGCTTTTACGAAAAAGCG	3660
Db	3601	AAGATCTAGTTTGGTGAANAATGACATTTAAGGAAAGTTCTGCTTTTACGAAAAAGCG	3660
QY	3661	TCCGAAAGAGAGACCTTAGCAGAGAGTCTAGCCCTTTCACCCATACACATTTGGCTAGG	3720
Db	3661	TCCGAAAGAGAGACCTTAGCAGAGAGTCTAGCCCTTTCACCCATACACATTTGGCTAGG	3720
QY	3721	GTTACCGAAGAGGGGCCAAGAAATTTAGAGTCTCAGAGAAGAACTTTCATAGTAGAGATG	3780
Db	3721	GTTACCGAAGAGGGGCCAAGAAATTTAGAGTCTCAGAGAAGAACTTTCATAGTAGAGATG	3780
QY	3781	AAGAGCTTCCCTGCTTCCAAACACTTGTATTATTGGTAAAGTAAACAATATACCTTCTCAGT	3840
Db	3781	AAGAGCTTCCCTGCTTCCAAACACTTGTATTATTGGTAAAGTAAACAATATACCTTCTCAGT	3840
QY	3841	CTACTAGGCAATAGCACCGTGTCTATCCAGATGTCTGTATAGAACACAGAGAGCAATTTAT	3900
Db	3841	CTACTAGGCAATAGCACCGTGTCTATCCAGATGTCTGTATAGAACACAGAGAGCAATTTAT	3900
QY	3901	TATCATTTGAAGAATAGCTTAAATGATGACGTACGTAAACAGGTAATTTGGCAAAAGCACTTC	3960
Db	3901	TATCATTTGAAGAATAGCTTAAATGATGACGTACGTAAACAGGTAATTTGGCAAAAGCACTTC	3960
QY	3961	AGGAACATCAACTTAGTAGAGAAAAAATAATGTTCTCTAGACTTGTTTTCTTACACAGTCA	4020

Db	3961	AGGAACATCACCTTATGTGAGGAAACAAAATGTTCTGTAGCTGTGTTTCTTCCACAGTCA	4020
Oy	4021	GTGAATTTGGAAACATCTGACATGCAATTAACAAACACCAGATCCCTTCTTGATGGTCTT	4080
Db	4021	GTGAATTTGGAAACATCTGACATGCAATTAACAAACACCAGATCCCTTCTTGATGGTCTT	4080
Oy	4081	CCAAACAATGAGGCATGCTGGAAGGCAGGAGATTGCTGAGTACAGAAATTTGG	4140
Db	4081	CCAAACAATGAGGCATGCTGGAAGGCAGGAGATTGCTGAGTACAGAAATTTGG	4140
Oy	4141	TTTCAGATGATGTAACAAGAGGAAACGGGCTTGGAAAGAAATATATACAGAGCAACCA	4200
Db	4141	TTTCAGATGATGTAACAAGAGGAAACGGGCTTGGAAAGAAATATATACAGAGCAACCA	4200
Oy	4201	TGGATTTCAAACTTTAGTGTGAGCAGCATCTGGGTGTGAGAGTGAACAGGCGTCTGAG	4260
Db	4201	TGGATTTCAAACTTTAGTGTGAGCAGCATCTGGGTGTGAGAGTGAACAGGCGTCTGAG	4260
Oy	4261	ACTGCTCAGGGGCTATCCTCTGAGGTGACATTTTAAACCTCAGCAGAGGATATACATGC	4320
Db	4261	ACTGCTCAGGGGCTATCCTCTGAGGTGACATTTTAAACCTCAGCAGAGGATATACATGC	4320
Oy	4321	AACATTAACCTGTATAAGCTCCAGCAGGAAATGTGCTGAACCTAGAGCTGTTAGACAGC	4380
Db	4321	AACATTAACCTGTATAAGCTCCAGCAGGAAATGTGCTGAACCTAGAGCTGTTAGACAGC	4380
Oy	4381	ATGGAGGAGGAGCTTCTAACACCTCCTCCATATTAAGTACCTTCTGGCCCTTGAG	4440
Db	4381	ATGGAGGAGGAGCTTCTAACACCTCCTCCATATTAAGTACCTTCTGGCCCTTGAG	4440
Oy	4441	ACCTGCGAAATCCAGAACAAAGCACATCAGAAAAAGCAGTATTAACTTCACAGAAAGTA	4500
Db	4441	ACCTGCGAAATCCAGAACAAAGCACATCAGAAAAAGCAGTATTAACTTCACAGAAAGTA	4500
Oy	4501	GTGAATACCCATATAAGCCAGAAATCCAGAAAGGCTTTCTGTGACAAATTTAGTGTCTG	4560
Db	4501	GTGAATACCCATATAAGCCAGAAATCCAGAAAGGCTTTCTGTGACAAATTTAGTGTCTG	4560
Oy	4561	CAGATAGTTCATCCAGTAATAAATAAGAACCAAGAGGTGAAAGGTCATCCCTCTTAAT	4620
Db	4561	CAGATAGTTCATCCAGTAATAAATAAGAACCAAGAGGTGAAAGGTCATCCCTCTTAAT	4620
Oy	4621	GCCCATCATTTAGATATAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG	4680
Db	4621	GCCCATCATTTAGATATAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG	4680
Oy	4681	ACTACCCATCTCAGAGAGAGCTCATTAAGTTGTTGATGTGAGAGGACAACGCTGAG	4740
Db	4681	ACTACCCATCTCAGAGAGAGCTCATTAAGTTGTTGATGTGAGAGGACAACGCTGAG	4740
Oy	4741	AGTCTGGGACACAGCATTTGACAGGAAACATCTTACTTGGCAAAGGAAACATCTAGAGGAA	4800
Db	4741	AGTCTGGGACACAGCATTTGACAGGAAACATCTTACTTGGCAAAGGAAACATCTAGAGGAA	4800
Oy	4801	CCCTTTACCTGGAATCTGGAATCACCCCTTCTCTGATGAGACCTCAATCTGATCCCTTGTG	4860
Db	4801	CCCTTTACCTGGAATCTGGAATCACCCCTTCTCTGATGAGACCTCAATCTGATCCCTTGTG	4860
Oy	4861	AAGACAGAGCCCCAGAGTCAGCTGTTGTTGCAACATACCATCTTCAACCTCTGCATTTGA	4920
Db	4861	AAGACAGAGCCCCAGAGTCAGCTGTTGTTGCAACATACCATCTTCAACCTCTGCATTTGA	4920
Oy	4921	AAATTTCCCAATTTGAAGTTGACAGATCTGCCAGAGTCAGGAGGATCTGATCTACTG	4980
Db	4921	AAATTTCCCAATTTGAAGTTGACAGATCTGCCAGAGTCAGGAGGATCTGATCTACTG	4980
Oy	4981	ATATGCTGGGTATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	5040
Db	4981	ATATGCTGGGTATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	5040
Oy	5041	CTTCAACGAAAGGGTCAACAAAGAAATGTCATAGTGTGCTGTGAGGCTGAGCCCAAG	5100
Db	5041	CTTCAACGAAAGGGTCAACAAAGAAATGTCATAGTGTGCTGTGAGGCTGAGCCCAAG	5100



QY	421	ATGCAAAACGCTATATATTTTGCAAAAAGAAATAACTCTCTCTTAACATCTTAAAGATG	480
Db	421	ATGCAAAACGCTATATATTTTGCAAAAAGAAATAACTCTCTTAACATCTTAAAGATG	480
QY	481	AAGTTTCTATATCCAAAGATATGGGCTCAGAAACCGTGCAGAAAGACTTCTACAGAGTG	540
Db	481	AAGTTTCTATATCCAAAGATATGGGCTCAGAAACCGTGCAGAAAGACTTCTACAGAGTG	540
QY	541	AACCCGAAAAATCCTTCTCTTGACAGAAACCACTTCAGTGTCCAACTCTTAACTTGGAA	600
Db	541	AACCCGAAAAATCCTTCTCTTGACAGAAACCACTTCAGTGTCCAACTCTTAACTTGGAA	600
QY	601	CTGTGAGAACTCTGAGGACAAAGCAGCGGGATATCAACCTCAAAAGACGTCTCATTTG	660
Db	601	CTGTGAGAACTCTGAGGACAAAGCAGCGGGATATCAACCTCAAAAGACGTCTCATTTG	660
QY	661	AATTGGAGATCGATTTCTCTGAGATACCGTTAATPAAGCAACTAATTGCAGGTGGGAG	720
Db	661	AATTGGAGATCGATTTCTCTGAGATACCGTTAATPAAGCAACTAATTGCAGGTGGGAG	720
QY	721	ATCAAGAAATTTGTACAAATCAACCCCTCAAGGACCAAGGATGAATTCAGTTTGGATTCG	780
Db	721	ATCAAGAAATTTGTACAAATCAACCCCTCAAGGACCAAGGATGAATTCAGTTTGGATTCG	780
QY	781	CAAAAAGGCGTCTGTGAATTTTCTGAGACGAGTATACAAATCTGAACTATCAAC	840
Db	781	CAAAAAGGCGTCTGTGAATTTTCTGAGACGAGTATACAAATCTGAACTATCAAC	840
QY	841	CCAGTAATAATGATTTGAAACCCACTGGAAGCCGTGCACCTGAGAGGCATCCAGAAAGT	900
Db	841	CCAGTAATAATGATTTGAAACCCACTGGAAGCCGTGCACCTGAGAGGCATCCAGAAAGT	900
QY	901	ATCAGGATAGTTCTGTTTCAAACTTCGATGTGGACCCATGTGGCCAAATATCTATGCCA	960
Db	901	ATCAGGATAGTTCTGTTTCAAACTTCGATGTGGACCCATGTGGCCAAATATCTATGCCA	960
QY	961	GCTCATTAACAGATGAGAAACAGCAGTTTATTACTACTTAAGACAGAAATGAATGTAGAA	1020
Db	961	GCTCATTAACAGATGAGAAACAGCAGTTTATTACTACTTAAGACAGAAATGAATGTAGAA	1020
QY	1021	AGCGTGAATTTCTGTATATAAAGCAAAACAGCCGTGTAGCAAGGCCAATTAACAGAT	1080
Db	1021	AGCGTGAATTTCTGTATATAAAGCAAAACAGCCGTGTAGCAAGGCCAATTAACAGAT	1080
QY	1081	GGCGCTGAAGTAAGGAAACATGTATGATGATGAGGCGGATCCGACGACAGAAAAAAGGTAG	1140
Db	1081	GGCGCTGAAGTAAGGAAACATGTATGATGATGAGGCGGATCCGACGACAGAAAAAAGGTAG	1140
QY	1141	ATCTGAATGCTGATCCCTGTGTGTGAGAGAAAAAGAAATGGAATAGCAGAACTCCATCT	1200
Db	1141	ATCTGAATGCTGATCCCTGTGTGTGAGAGAAAAAGAAATGGAATAGCAGAACTCCATCT	1200
QY	1201	CAGAAATCCTAGAGATACTGAAGTGTTCCTTGATTAACACTTAATAGCAGCATTCAGA	1260
Db	1201	CAGAAATCCTAGAGATACTGAAGTGTTCCTTGATTAACACTTAATAGCAGCATTCAGA	1260
QY	1261	AAGTTAATGAGTGTGTTTCCAGAACTGATGAACCTTATGAGTTCTGATACCTACATGATG	1320
Db	1261	AAGTTAATGAGTGTGTTTCCAGAACTGATGAACCTTATGAGTTCTGATACCTACATGATG	1320
QY	1321	GGGAGTCTGAATCAAAATGCCAAAGTAGTGTATTTGAGCGTTCTTAATGAGGTAGATG	1380
Db	1321	GGGAGTCTGAATCAAAATGCCAAAGTAGTGTATTTGAGCGTTCTTAATGAGGTAGATG	1380
QY	1381	AATATTTCTGGTTCTTCAGAGAAATATAGCTTACTGGCCAGATGATCTCATAGGCTTTAA	1440
Db	1381	AATATTTCTGGTTCTTCAGAGAAATATAGCTTACTGGCCAGATGATCTCATAGGCTTTAA	1440
QY	1441	TATGTAAAGTGAAGAGTGTACATCCCAATAGTAGAGAGATTAATTTGAGGCAAAATAT	1500
Db	1441	TATGTAAAGTGAAGAGTGTACATCCCAATAGTAGAGAGATTAATTTGAGGCAAAATAT	1500
QY	1501	TTGGGAAAAACCTATGGAAGAAGCCAGCTCCCAACTTAAGCATATGATACTGAAATTC	1560

Db	1501	TTGGGAAAACCTATCGAAGAGCAAGCCCTCCCACTTAAGCCATGTAACTGAAAAATC	1560
QY	1561	TAATATATAGGAGCATTTGTTACTGTAGCCACAGATATATCAAGAGAGTCCCTCCACAAATA	1620
Db	1561	TAATATATAGAGACATTTTGTACTGAGCCACAGATATATACAGAGAGTCCCTCCACAAATA	1620
QY	1621	AATTAAAGCGTAAAGAGAGACCTACATCAGGCCCTTATCTGAGGATTTTATCAAGAAAG	1680
Db	1621	AATTAAAGCGTAAAGAGAGACCTACATCAGGCCCTTATCTGAGGATTTTATCAAGAAAG	1680
QY	1661	CAGATTGTGGCACTTCAAAAAGATCCCTGAAATGATTAATCAAGGAACTAACCAAGCGAGC	1740
Db	1661	CAGATTGTGGCACTTCAAAAAGATCCCTGAAATGATTAATCAAGGAACTAACCAAGCGAGC	1740
QY	1741	AGAATGTGCAAGTGTATGATATATTTACTATATAGTGTCATAGAGATTAACCAAAAGSTGAT	1800
Db	1741	AGAATGTGCAAGTGTATGATATATTTACTATATAGTGTCATAGAGATTAACCAAAAGSTGAT	1800
QY	1801	CTATTTCAGAAATGAGAAAAATCCCTAACCCAAATGAATCACTCGAAAAAGAAATCTGCTTTCA	1860
Db	1801	CTATTTCAGAAATGAGAAAAATCCCTAACCCAAATGAATCACTCGAAAAAGAAATCTGCTTTCA	1860
QY	1861	AAAGCAAAAGCTGAACTTATTAAGCAGCAGTATATAGCAATATGGAATCGAATTAATATTC	1920
Db	1861	AAAGCAAAAGCTGAACTTATTAAGCAGCAGTATATAGCAATATGGAATCGAATTAATATTC	1920
QY	1921	ACAATTTCAAAAAGCCCTTAATAAAGATATAGCGTGAGGAGGAAAGCTCTTCCACGAGCATATTC	1980
Db	1921	ACAATTTCAAAAAGCCCTTAATAAAGATATAGCGTGAGGAGGAAAGCTCTTCCACGAGCATATTC	1980
QY	1981	ATGCGCTTGAACCTAGTACTAGTATGAAATCTTAAGGCCCACTAATTTGTACTCAATTTGCAGAA	2040
Db	1981	ATGCGCTTGAACCTAGTACTAGTATGAAATCTTAAGGCCCACTAATTTGTACTCAATTTGCAGAA	2040
QY	2041	TTGATAGTTGTTCTAGCAGTGAAGATTAAGAAAAAAAGTACACCAATATGCCAGTCA	2100
Db	2041	TTGATAGTTGTTCTAGCAGTGAAGATTAAGAAAAAAAGTACACCAATATGCCAGTCA	2100
QY	2101	GGCAGAGAGAAACCTACAACTCACTGAAAGGTTAAAGAAACCTGCAACTGAGACCAAGAGA	2160
Db	2101	GGCAGAGAGAAACCTACAACTCACTGAAAGGTTAAAGAAACCTGCAACTGAGACCAAGAGA	2160
QY	2161	GTAACAAACCCAAATGAAAGACAGATTAAGACATGACAGTAGTACTTTCCAGAGCTGA	2220
Db	2161	GTAACAAACCCAAATGAAAGACAGATTAAGACATGACAGTAGTACTTTCCAGAGCTGA	2220
QY	2221	AGTTAACAAATGACACCTGCTTTTACTAAGTGTTCAATATCCAGTAGTAACCTTAAGAAT	2280
Db	2221	AGTTAACAAATGACACCTGCTTTTACTAAGTGTTCAATATCCAGTAGTAACCTTAAGAAT	2280
QY	2281	TTTGTCAAATCCTAAGCCTTCCCAAGAGAGAAAAAGAAAGAAACCTGAACACTTAAAGTGT	2340
Db	2281	TTTGTCAAATCCTAAGCCTTCCCAAGAGAGAAAAAGAAAGAAACCTGAACACTTAAAGTGT	2340
QY	2341	CTAATTAATGCTGAAGAGACCCCAAGATCTCATGTATTAAGTGAGGAAAGGGTTTTCGAAACTG	2400
Db	2341	CTAATTAATGCTGAAGAGACCCCAAGATCTCATGTATTAAGTGAGGAAAGGGTTTTCGAAACTG	2400
QY	2401	AAAGATCTGTAGAGAGTAGCAGTATTTTCATTTGTAACCTGTAAGTATATGATATGGAACCTCAGG	2460
Db	2401	AAAGATCTGTAGAGAGTAGCAGTATTTTCATTTGTAACCTGTAAGTATATGATATGGAACCTCAGG	2460
QY	2461	AAAGTATCTGCTTACTGGAAGTATGACACTCTAAGGGAAGGCAAAAAACAGAACCAATATAT	2520
Db	2461	AAAGTATCTGCTTACTGGAAGTATGACACTCTAAGGGAAGGCAAAAAACAGAACCAATATAT	2520
QY	2521	GTTGAGAGTCACTGAGACATTTTGAAGAAACCCCAAGGGGACTAATTCATGCGTGTTCGCAAG	2580
Db	2521	GTTGAGAGTCACTGAGACATTTTGAAGAAACCCCAAGGGGACTAATTCATGCGTGTTCGCAAG	2580
QY	2581	ATATATAGAAATGACACAGAAAGGCTTTAAGTATCCATTTGGGACATGAAGTTAACCACAGTCT	2640



2581 ATATAGAAATGACACAGAAAGCCTTTAAGATATCCATTGGACATGAAATTACCCACAGTC 2640  
2641 GGGAAACACGATGAAATGAGAAAGTGAATGCTCAGTATTTGAGAAATCAT 2700  
2641 GGGAAACACGATGAAATGAGAAAGTGAATGCTCAGTATTTGAGAAATCAT 2700  
2701 TCAAGGTTTCAAGCGCCAGCTCATTTGCTGTTTTCAAATCCAGAAATGAGAGAG 2760  
2701 TCAAGGTTTCAAGCGCCAGCTCATTTGCTGTTTTCAAATCCAGAAATGAGAGAG 2760  
2761 AATGTGCAACATTTCTGCCCCACTCTGGGTCCTTAAAGAAACAAAGTCACAAATGCT 2820  
2761 AATGTGCAACATTTCTGCCCCACTCTGGGTCCTTAAAGAAACAAAGTCACAAATGCT 2820  
2821 TTGAATGTGAACAAAGAAAGAAATATCAGAAAGAAATGAGTCAATATCAAGCTGTAC 2880  
2821 TTGAATGTGAACAAAGAAAGAAATATCAGAAAGAAATGAGTCAATATCAAGCTGTAC 2880  
2881 AGACAGTTAATATCACTGACAGGCTTTCTGTTGGTTCAGAAAGATTAAGCCAGTTGATA 2940  
2881 AGACAGTTAATATCACTGACAGGCTTTCTGTTGGTTCAGAAAGATTAAGCCAGTTGATA 2940  
2941 ATGCCAAATGATGATCAAGAGAGGCTCTAGGTTTGTCTATCATCTCAGTTGAGAGCA 3000  
2941 ATGCCAAATGATGATCAAGAGAGGCTCTAGGTTTGTCTATCATCTCAGTTGAGAGCA 3000  
3001 ACGAACCTGACATCTACTTCAAAATTAACATGACATTTTACAAACCCATATCTGTATAC 3060  
3001 ACGAACCTGACATCTACTTCAAAATTAACATGACATTTTACAAACCCATATCTGTATAC 3060  
3061 CACCACTTTTCCATCAAGCTCATTTGTAAACTAAATGTAAGAAATCTGCTAGAG 3120  
3061 CACCACTTTTCCATCAAGCTCATTTGTAAACTAAATGTAAGAAATCTGCTAGAG 3120  
3061 CACCACTTTTCCATCAAGCTCATTTGTAAACTAAATGTAAGAAATCTGCTAGAG 3120  
3121 AAACTTTGAGAAACATTCATGTCACCTGAAAGAAATGGGAAATGAGAACCTTCCAA 3180  
3121 AAACTTTGAGAAACATTCATGTCACCTGAAAGAAATGGGAAATGAGAACCTTCCAA 3180  
3181 GTACAGTGAACATTAAGCCGTAATTAACATTAAGAAATGTTTAAAGAACCCAGCT 3240  
3181 GTACAGTGAACATTAAGCCGTAATTAACATTAAGAAATGTTTAAAGAACCCAGCT 3240  
3241 CAAGCAATTAATGAAGTGAAGTTCAGTCAATGAAGTGGGCTCCAGTATTAATGAA 3300  
3241 CAAGCAATTAATGAAGTGAAGTTCAGTCAATGAAGTGGGCTCCAGTATTAATGAA 3300  
3301 TAGCTTCAGTGAAGAAATTCACAGCAAGTATGAAACAGAGGCGCAAAATGTA 3360  
3301 TAGCTTCAGTGAAGAAATTCACAGCAAGTATGAAACAGAGGCGCAAAATGTA 3360  
3361 ATGCTATGCTTAGATTAGGGGTTTTCACACCTGAGTCTATTAACAAGTCTTCTGGA 3420  
3361 ATGCTATGCTTAGATTAGGGGTTTTCACACCTGAGTCTATTAACAAGTCTTCTGGA 3420  
3421 GTATATGTGAAGCTCTGGAATTAAGAAAGCAAGATTAAGAAAGTTCAGACTGTA 3480  
3421 GTATATGTGAAGCTCTGGAATTAAGAAAGCAAGATTAAGAAAGTTCAGACTGTA 3480  
3481 ATACAGTTTCTGTCATATCTGATTTCAATTAAGTGAAGCAAGCTATGGAAGTGTGTC 3540  
3481 ATACAGTTTCTGTCATATCTGATTTCAATTAAGTGAAGCAAGCTATGGAAGTGTGTC 3540  
3541 ATGCATCTCAGGTTTGTCTGAGACACCTGATGACCTGTAGATGATGGAATTAAGG 3600  
3541 ATGCATCTCAGGTTTGTCTGAGACACCTGATGACCTGTAGATGATGGAATTAAGG 3600  
3601 AAGATCTAGTTTGTCTGGAATTAAGCAATTAAGAAAGTCTGCTGTTTAAAGAAAGCG 3660  
3601 AAGATCTAGTTTGTCTGGAATTAAGCAATTAAGAAAGTCTGCTGTTTAAAGAAAGCG 3660  
3661 TCCAGAAAGAGAGCTTAGCAGAGTCTAGCCCTTACCCATTAACATTTGGCTAGG 3720  
3661 TCCAGAAAGAGAGCTTAGCAGAGTCTAGCCCTTACCCATTAACATTTGGCTAGG 3720

3721 GTTACCGAAGAGGGGCCAAGAAATTAAGTGTCTCAGAAAGAGAACTTATCTAGAGAGATG 3780  
3721 GTTACCGAAGAGGGGCCAAGAAATTAAGTGTCTCAGAAAGAGAACTTATCTAGAGAGATG 3780  
3781 AAGAGCTTCCCTGCTTCAACACTTGTATTTTGTAAAGTAAACAATATACCTTCTCACT 3840  
3781 AAGAGCTTCCCTGCTTCAACACTTGTATTTTGTAAAGTAAACAATATACCTTCTCACT 3840  
3841 CTACTAGGCAATAGCAGCGTGTCTACCGAGTGTCTGTCTAAGAAACACAGAGGAAATTTAT 3900  
3841 CTACTAGGCAATAGCAGCGTGTCTACCGAGTGTCTGTCTAAGAAACACAGAGGAAATTTAT 3900  
3901 TATCATTAAGAAATAGCTTAAATGACGAGTCAACAGGAAATATGGAAGAGGATCTTC 3960  
3901 TATCATTAAGAAATAGCTTAAATGACGAGTCAACAGGAAATATGGAAGAGGATCTTC 3960  
3961 AGGAACATCACTTAGTGAAGAAACAAATGTTCTGCTAGCTGTTTCTTCAAGTGA 4020  
3961 AGGAACATCACTTAGTGAAGAAACAAATGTTCTGCTAGCTGTTTCTTCAAGTGA 4020  
4021 GTGAATTTGAAGAGCTTACGCAAAATTAACAAACCCAGAGATCCTTCTGATTTGCTT 4080  
4021 GTGAATTTGAAGAGCTTACGCAAAATTAACAAACCCAGAGATCCTTCTGATTTGCTT 4080  
4081 CCAAAACAAATGAGGATCACTGTAAGAGCCAGGAGTTGGTCTGAGTGAAGAAATTTGG 4140  
4081 CCAAAACAAATGAGGATCACTGTAAGAGCCAGGAGTTGGTCTGAGTGAAGAAATTTGG 4140  
4141 TTTGATGATGATGAAGAAAGAGAGCGGCTTGAAGAAATTAATCAAGAAAGCAAGCA 4200  
4141 TTTGATGATGATGAAGAAAGAGAGCGGCTTGAAGAAATTAATCAAGAAAGCAAGCA 4200  
4201 TGGATTTAACTTAAGTGAAGAGAGCACTGAGGTTGAGTGAAGTGAAGTGAAGTGAAG 4260  
4201 TGGATTTAACTTAAGTGAAGAGAGCACTGAGGTTGAGTGAAGTGAAGTGAAGTGAAG 4260  
4261 ACTGCTCAGGAGCTATCTCTCAGAGTGAATTTTAAACCACTGACAGAGAGGATTAACATTC 4320  
4261 ACTGCTCAGGAGCTATCTCTCAGAGTGAATTTTAAACCACTGACAGAGAGGATTAACATTC 4320  
4321 AACATTAACCTGATTAAGAGTCAACAGAGAAATGAGTCAACTAGAGCTGCTTGAAGACG 4380  
4321 AACATTAACCTGATTAAGAGTCAACAGAGAAATGAGTCAACTAGAGCTGCTTGAAGACG 4380  
4381 ATGGAGGCCAGGCTTCTTAACAGCTTCAACCTTCAATTAAGTGAAGTGAAGTGAAG 4440  
4381 ATGGAGGCCAGGCTTCTTAACAGCTTCAACCTTCAATTAAGTGAAGTGAAGTGAAG 4440  
4441 ACCGCGAATTCAGAAACAAAGCAATCAAGAAACAGTATTAATCTTCAAGAAAGTGA 4500  
4441 ACCGCGAATTCAGAAACAAAGCAATCAAGAAACAGTATTAATCTTCAAGAAAGTGA 4500  
4501 GTGAATTAACCTTATTAAGCAGATTCAGAAAGGCTTCTCTGCAAGAGTGAAGTGAAG 4560  
4501 GTGAATTAACCTTATTAAGCAGATTCAGAAAGGCTTCTCTGCAAGAGTGAAGTGAAG 4560  
4561 CAGATAGTCTTACAGTAAATAAGAAACAGAGAGTGAAGAGTCACTCCCTTCAAT 4620  
4561 CAGATAGTCTTACAGTAAATAAGAAACAGAGAGTGAAGAGTCACTCCCTTCAAT 4620  
4621 GCCCATCATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 4680  
4621 GCCCATCATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 4680  
4681 ACTTACCATTAAGTGAAGAGAGTCAATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 4740  
4681 ACTTACCATTAAGTGAAGAGAGTCAATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 4740  
4741 AGTCTGGGCAACAGATTTAGAGGAAACATCTTAAGTGAAGTGAAGTGAAGTGAAGTGAAG 4800  
4741 AGTCTGGGCAACAGATTTAGAGGAAACATCTTAAGTGAAGTGAAGTGAAGTGAAGTGAAG 4800



4801 CCCCTTACCTGGAATCTGGAATCAGCTCTTCTGTATGACCTGGAATCTGATCTCTG 4860  
 4801 CCCCTTACCTGGAATCTGGAATCAGCTCTTCTGTATGACCTGGAATCTGATCTCTG 4860  
 4861 AAGACAGAGCCCGAGCTGAGCTGCTGTTGGCAACATACATCTTCAACCTGCAATGA 4920  
 4861 AAGACAGAGCCCGAGCTGAGCTGCTGTTGGCAACATACATCTTCAACCTGCAATGA 4920  
 4861 AAGACAGAGCCCGAGCTGAGCTGCTGTTGGCAACATACATCTTCAACCTGCAATGA 4920  
 4921 AAGTCCCAATGGAAGTTGCAGAAATCTGCCAGAGTCCAGCTGCTCATCTACTG 4980  
 4921 AAGTCCCAATGGAAGTTGCAGAAATCTGCCAGAGTCCAGCTGCTCATCTACTG 4980  
 4981 ATACTGCTGGGTATATGCAATGGAAGAAAGTGTACAGAGGAGAGCCAGAAATGACAG 5040  
 4981 ATACTGCTGGGTATATGCAATGGAAGAAAGTGTACAGAGGAGAGCCAGAAATGACAG 5040  
 5041 CTTCACAGAAAGGGTCAACAAAGAAATGTCATGAGTGTGCTGAGCCCTGACCCCAAG 5100  
 5041 CTTCACAGAAAGGGTCAACAAAGAAATGTCATGAGTGTGCTGAGCCCTGACCCCAAG 5100  
 5101 AATTTATGCTGCTAGCAAGTTGCCAGAAACACCATCATCTTAATCTAATTA 5160  
 5101 AATTTATGCTGCTAGCAAGTTGCCAGAAACACCATCATCTTAATCTAATTA 5160  
 5161 CTGAGAGAGACTCATCATGTTGTTATGAAACAGATGCTGATGTTGTGTGAGACGACAC 5220  
 5161 CTGAGAGAGACTCATCATGTTGTTATGAAACAGATGCTGATGTTGTGTGAGACGACAC 5220  
 5221 TGAATATTTTCTAGGAATTCGGGAGGAAATGGGATAGTATTTTCTGGGTGACCC 5280  
 5221 TGAATATTTTCTAGGAATTCGGGAGGAAATGGGATAGTATTTTCTGGGTGACCC 5280  
 5281 AGTCTATTTAAAGAGAAAGAAATGCTGAATGAGATGTTTGAAGTCAAGAGAGATGTCG 5340  
 5281 AGTCTATTTAAAGAGAAAGAAATGCTGAATGAGATGTTTGAAGTCAAGAGAGATGTCG 5340  
 5341 TCAATGGAAGAAACCCAGAGTCCAAAGCGAGCAAGAAATCCAGAGAGAGAAAGATCT 5400  
 5341 TCAATGGAAGAAACCCAGAGTCCAAAGCGAGCAAGAAATCCAGAGAGAGAAAGATCT 5400  
 5401 TCAGGGGGCTGAATGCTGTTGCTATGGCCCTTCACCAATGCCAGATCACTG 5460  
 5401 TCAGGGGGCTGAATGCTGTTGCTATGGCCCTTCACCAATGCCAGATCACTG 5460  
 5461 AATGATGATGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5520  
 5461 AATGATGATGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5520  
 5521 GCACAGGTGTCACCCCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5580  
 5521 GCACAGGTGTCACCCCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5580  
 5581 TCACATGCAATGGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5640  
 5581 TCACATGCAATGGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5640  
 5641 GTGTAGACACTCTACAGTGCAGAGAGTGCAGACCTGATGCTGCTGCTGCTGCTGCTGCTG 5700  
 5641 GTGTAGACACTCTACAGTGCAGAGAGTGCAGACCTGATGCTGCTGCTGCTGCTGCTGCTG 5700  
 5701 GCCACTACTGA 5711  
 5701 GCCACTACTGA 5711

RESULT 15  
 AAT17438  
 ID AAT17438 standard; cDNA: 5914 BP.  
 XX AAT17438;  
 XX AAT17438;  
 DT 30-SEP-1996 (first entry)  
 XX

DE BRCA1 coding sequence.  
 XX Cancer therapy; breast and ovarian cancer predisposing gene; immunogen;  
 KW antibody production; germline alteration; probe; lesion neoplasia; human;  
 KW gene therapy; protein replacement therapy; protein mimetic; BRCA1; ds.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT CDS 120..5711  
 FT /tag= a  
 FT /product= BRCA1  
 XX MO9605306-A2.  
 XX 22-FEB-1996.  
 XX 11-AUG-1995; 95WO-US10202.  
 XX 07-JUN-1995; 95US-0483553.  
 XX 12-AUG-1994; 94US-0289221.  
 XX 02-SEP-1994; 94US-0300266.  
 XX 16-SEP-1994; 94US-0308104.  
 XX 29-NOV-1994; 94US-0348824.  
 XX 24-MAR-1995; 95US-0409305.  
 XX 07-JUN-1995; 95US-0480784.  
 PA (MYRI-) MYRIAD GENETICS INC.  
 PA (CAN-) CANCER INST.  
 PA (RECH-) CENT RECH DU CHUL.  
 PI Durocher F, Emi M, Nakamura Y, Simard J, Shattuck-Eidens DM.  
 XX WPI: 1996-139702/14.  
 XX P-PSDB: MAR81481.  
 XX New nucleic acid and polypeptide for mutant or polymorphic BRCA1  
 PT gene - for diagnosis and therapy of human breast and ovarian cancer  
 PT and for diagnosing pre-disposition to these cancers  
 XX  
 XX Claim 1; Page 119-128; 218pp; English.  
 XX  
 XX This sequence represents the cDNA of the human breast and ovarian cancer  
 CC predisposing gene (BRCA1). Proteins encoded by mutations of this  
 CC sequence (see AAT17439-T17453 and AAT17455-T17529) can be used as  
 CC immunogens for antibody production. An additional BRCA1 mutant, with a T  
 CC to G mutation and 59 nucleotide insert (encoding a protein which  
 CC terminates at residue 75) is also claimed, however the exact location of  
 CC the mutations is not specified so the sequence could not be created. The  
 CC mutant BRCA1 genes have at least 1 mutation or polymorphism in comparison  
 CC to this sequence. By detecting a germline alteration in this gene, a  
 CC predisposition for breast and ovarian cancer can be diagnosed. In one  
 CC method, BRCA1 mRNA isolated from a tissue sample from a subject has a  
 CC probe, corresponding to a fragment of this sequence (or an  
 CC allele-specific probe for a mutation of this sequence), added to it.  
 CC The conditions allow for hybridisation of the probe to the mRNA, and any  
 CC hybridisation which occurs is detected. Alternatively the BRCA1 gene in  
 CC the tissue sample is isolated, and a shift in electrophoretic mobility  
 CC of single stranded DNA from the sample on a non-denaturing  
 CC polyacrylamide gel indicates a mutation. These methods of detection can  
 CC also diagnose a lesion neoplasia associated with the BRCA1 locus. The  
 CC methods may be used in gene therapy, protein replacement therapy and  
 CC protein mimetics, and may be used to screen for drugs in cancer therapy.  
 XX  
 SO Sequence 5914 BP; 2006 A; 1156 C; 1316 G; 1436 T; 0 other;  
 Query Match 99.9%; Score 5707.8; DB 17; Length 5914;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 5709; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 OY 1 AGCTGCTGAGAGACTTCTGAGAGCCCGACCAAGGCTGTGGGTTCTCAGATACTGGGCC 60  
 DB 1 AGCTGCTGAGAGACTTCTGAGAGCCCGACCAAGGCTGTGGGTTCTCAGATACTGGGCC 60

QY	61	CCGCGCTCAGGAGGCCCTTACACCCCTCTGCTCGGGTAAAGTTCAATGGACAGAAAGAA	120
Db	61	CTGCGCTCAGGAGGCCCTTACACCCCTCTGCTCGGGTAAAGTTCAATGGACAGAAAGAA	120
QY	121	TGGATTATTCGCTCTCTGCGCTTGAAGAATGCAAAATGTCATTATATGTCAGAAAA	180
Db	121	TGGATTATTCGCTCTCTGCGCTTGAAGAATGCAAAATGTCATTATATGTCAGAAAA	180
QY	181	TCTTGAAGTGTCCCATCTGCTGTGAGTTGATCAAGCACTGTGTCCCAAGGTGACC	240
Db	181	TCTTGAAGTGTCCCATCTGCTGTGAGTTGATCAAGCACTGTGTCCCAAGGTGACC	240
QY	241	ACATATTTTGGCAAAATTTTGCATGTCGAAACTTCTCACAGAGAAAGGCCCTTACAGT	300
Db	241	ACATATTTTGGCAAAATTTTGCATGTCGAAACTTCTCACAGAGAAAGGCCCTTACAGT	300
QY	301	GTCCTTTATGTAGAAATGATATTAACCAAAAGAGCCTCAGAAAGATGCGAGATTGTGTC	360
Db	301	GTCCTTTATGTAGAAATGATATTAACCAAAAGAGCCTCAGAAAGATGCGAGATTGTGTC	360
QY	361	AACTGTTTGAAGAGCTATTGAAATCATTTGTGCTTTTACGTTTGACAGAGTTTGGAGT	420
Db	361	AACTGTTTGAAGAGCTATTGAAATCATTTGTGCTTTTACGTTTGACAGAGTTTGGAGT	420
QY	421	ATGCAACAGAGCTATATTTTGCAGAAAAAGAAAAATACCTCTCTCGAACATCTAAAGATG	480
Db	421	ATGCAACAGAGCTATATTTTGCAGAAAAAGAAAAATACCTCTCTCGAACATCTAAAGATG	480
QY	481	AAGTTTCTATCATCCAAAGTATGGGCTACAGAAAACCGTCCAAAAGACTTCTACAGAGT	540
Db	481	AAGTTTCTATCATCCAAAGTATGGGCTACAGAAAACCGTCCAAAAGACTTCTACAGAGT	540
QY	541	AAACCCGAAAATTCCTTCTTGCAGAGAACCAAGTCTCAGTGTCCAACTCTTAACCTTGGAA	600
Db	541	AAACCCGAAAATTCCTTCTTGCAGAGAACCAAGTCTCAGTGTCCAACTCTTAACCTTGGAA	600
QY	601	CTGTGAGAACTCTGAGCAAAAGCAGCGATACAACTCAAAAGACGCTGTCTACACTG	660
Db	601	CTGTGAGAACTCTGAGCAAAAGCAGCGATACAACTCAAAAGACGCTGTCTACACTG	660
QY	661	AATTGGAGTCTGATTCCTTCTGAGAGATCCCTTAATAAGCAACTTATTTGCACTGTGGAG	720
Db	661	AATTGGAGTCTGATTCCTTCTGAGAGATCCCTTAATAAGCAACTTATTTGCACTGTGGAG	720
QY	721	ATCAGAGATTTGTTACAAATCACCCCTCAGAGAACCGAGGATGGAATCAGTTTGGATTCTG	780
Db	721	ATCAGAGATTTGTTACAAATCACCCCTCAGAGAACCGAGGATGGAATCAGTTTGGATTCTG	780
QY	781	CAAAAAAGGCTGCTTGTGAATTTTCTGAGAGAGATGTAAACAAATCTGAACATCATCAAC	840
Db	781	CAAAAAAGGCTGCTTGTGAATTTTCTGAGAGAGATGTAAACAAATCTGAACATCATCAAC	840
QY	841	CCAGTAAATTAAGATTGTAACACCACTGGAAGCGTGCAGAGCGTAAGAGCATCCAGAAAAAGT	900
Db	841	CCAGTAAATTAAGATTGTAACACCACTGGAAGCGTGCAGAGCGTAAGAGCATCCAGAAAAAGT	900
QY	901	ATCAGGAGTGTCTGTTTCAAACTTGCAATGAGAGCATGTGAGCAAAATACATCATGCCA	960
Db	901	ATCAGGAGTGTCTGTTTCAAACTTGCAATGAGAGCATGTGAGCAAAATACATCATGCCA	960
QY	961	GCTCATTTACAGCATGAGAACAGCAGTTTATTTACTACATAAAGACAGATGAATGTAGAA	1020
Db	961	GCTCATTTACAGCATGAGAACAGCAGTTTATTTACTACATAAAGACAGATGAATGTAGAA	1020
QY	1021	AGGCTGAATTTCTGTATATTAAGAACAAGCCTGCGTTAGCAAGAGACCTCAACATPACAGAT	1080
Db	1021	AGGCTGAATTTCTGTATATTAAGAACAAGCCTGCGTTAGCAAGAGACCTCAACATPACAGAT	1080
QY	1081	GGGCTGGAAGTAAAGAAACATGTATATAGGGGAGATCCGAGACAGAAAAAAGGTAG	1140
Db	1081	GGGCTGGAAGTAAAGAAACATGTATATAGGGGAGATCCGAGACAGAAAAAAGGTAG	1140
QY	1141	ATCTGAATGCTGATCCCTGTGTGAGAGAAAAAGATGAAATGACAGAAACTGTCCATGCT	1200
Db	1141	ATCTGAATGCTGATCCCTGTGTGAGAGAAAAAGATGAAATGACAGAAACTGTCCATGCT	1200
QY	1201	CAGAGATTCCTAGAGATCTGAAATGTCTCTTGATTAACACCTAATTAACAGATTTCAGA	1260
Db	1201	CAGAGATTCCTAGAGATCTGAAATGTCTCTTGATTAACACCTAATTAACAGATTTCAGA	1260
QY	1261	AAGTTAATGAGTGGTTTCCAGAGATGATGAAGCTGTAGTGTGTGATGATCATATGATG	1320
Db	1261	AAGTTAATGAGTGGTTTCCAGAGATGATGAAGCTGTAGTGTGTGATGATCATATGATG	1320
QY	1321	GGGAGTCTGAATCAAAATGCCAAGTACGTGATGTATTGCACTTCTTAATGAGTATG	1380
Db	1321	GGGAGTCTGAATCAAAATGCCAAGTACGTGATGTATTGCACTTCTTAATGAGTATG	1380
QY	1381	AATATTCTGTGTTCTTCAAGAAATAGACTACTGAGGCACTGATCTCATGAGGCTTTAA	1440
Db	1381	AATATTCTGTGTTCTTCAAGAAATAGACTACTGAGGCACTGATCTCATGAGGCTTTAA	1440
QY	1441	TATGTAAAAAGTAAAGATTCCTCAAAATCAGTAAAGATTAATTTGAAGCAAAATAT	1500
Db	1441	TATGTAAAAAGTAAAGATTCCTCAAAATCAGTAAAGATTAATTTGAAGCAAAATAT	1500
QY	1501	TTGGGAAAAACCTATCCGAGAAAGGCAAGCCTCCCAACTTAACCTGTATCTGAAATATC	1560
Db	1501	TTGGGAAAAACCTATCCGAGAAAGGCAAGCCTCCCAACTTAACCTGTATCTGAAATATC	1560
QY	1561	TAAATTATGAGAGCATTTTGTACTAGACACAGATTAATACAAAGAGCCTGCCACAAATA	1620
Db	1561	TAAATTATGAGAGCATTTTGTACTAGACACAGATTAATACAAAGAGCCTGCCACAAATA	1620
QY	1621	AATTAAGCGTAAAGAGACCTACATCAGAGCCTTCATCTCGAGATTTTATCAAGAAAG	1680
Db	1621	AATTAAGCGTAAAGAGACCTACATCAGAGCCTTCATCTCGAGATTTTATCAAGAAAG	1680
QY	1681	CAGATTGTGCGATTCAAAAGACCTCTGAATGTAATACAGGAATCAACCAAGGAGC	1740
Db	1681	CAGATTGTGCGATTCAAAAGACCTCTGAATGTAATACAGGAATCAACCAAGGAGC	1740
QY	1741	AGAATGGTCAAGTATGATTAATTAATAGTGTGATGAGAAATTAACCAAAAGCTGATTT	1800
Db	1741	AGAATGGTCAAGTATGATTAATTAATAGTGTGATGAGAAATTAACCAAAAGCTGATTT	1800
QY	1801	CTATTCAAAATGAGAAAAATCTTAACCATTAATACATCTCGAAAAAGAAATGCTTTCA	1860
Db	1801	CTATTCAAAATGAGAAAAATCTTAACCATTAATACATCTCGAAAAAGAAATGCTTTCA	1860
QY	1861	AAACGAAAGCTGAACCTATAAGACAGCTGTAAAGCAATATGGAACCTGAAATTAATATCC	1920
Db	1861	AAACGAAAGCTGAACCTATAAGACAGCTGTAAAGCAATATGGAACCTGAAATTAATATCC	1920
QY	1921	ACAATTTCAAAAAGCACTTAAGAAAGATAGGCTGAGGAGAACTTCTACAGGCAATATTC	1980
Db	1921	ACAATTTCAAAAAGCACTTAAGAAAGATAGGCTGAGGAGAACTTCTACAGGCAATATTC	1980
QY	1981	ATGGCTTGAACCTAGTGTGATGAGAAATCTAAGCCCACTAATGTCTGATTTGCAAA	2040</

Db	2221	AGTTAACAAATGCACTTGTTCTTTTCTTAAGTGTTCAAATACCAAGTAACCTTAAGAT	2280
OY	2281	TTGTCAATCCAGAGCTTCCAGAGAGAGAGAAAGAACTAGAAACACTTAAGTGT	2340
Db	2282	TTGTCAATCCAGAGCTTCCAGAGAGAGAGAAAGAAAGAACTAGAAACACTTAAGTGT	2340
OY	2341	CTAATTAATGCTGAAGAGCCCAAGACTCATGTTAAGTGAGAAAGGTTTGGCAACTG	2400
Db	2341	CTAATTAATGCTGAAGAGCCCAAGACTCATGTTAAGTGAGAAAGGTTTGGCAACTG	2400
OY	2401	AAAGTGTGTGAGAGTAGACAGTATTTTCATGTGTACCTGGTACATGATTTATGGCACTCAGG	2460
Db	2401	AAAGTGTGTGAGAGTAGACAGTATTTTCATGTGTACCTGGTACATGATTTATGGCACTCAGG	2460
OY	2461	AAAGTATCTCGTTTACTGGAAAGTTAGCACTCTAGGGAAGAGCAAAAACAGAACCAATAAAT	2520
Db	2461	AAAGTATCTCGTTTACTGGAAAGTTAGCACTCTAGGGAAGAGCAAAAACAGAACCAATAAAT	2520
OY	2521	GTGTGAGTCACTGTGCAAGACTTTGAAAAACCCCAAGGACTAATTCATGTGTTTCCAAG	2580
Db	2521	GTGTGAGTCACTGTGCAAGACTTTGAAAAACCCCAAGGACTAATTCATGTGTTTCCAAG	2580
OY	2581	ATATATAGAAATGACACAGAAAGGCTTTAAGTATCCATTGGGACATGAAGTTAACCAAGTC	2640
Db	2581	ATATATAGAAATGACACAGAAAGGCTTTAAGTATCCATTGGGACATGAAGTTAACCAAGTC	2640
OY	2641	GGGAACACAGCATAGAAATGGAAGAAAGTGAATCTGATGCTCAGTATTTGCGAATACAT	2700
Db	2641	GGGAACACAGCATAGAAATGGAAGAAAGTGAATCTGATGCTCAGTATTTGCGAATACAT	2700
OY	2701	TCAAGGTTTCAAGAGCGCCAGTCAATTTGCTCTGTGTTTCAAAATCCAGAAATGCCAAGAGG	2760
Db	2701	TCAAGGTTTCAAGAGCGCCAGTCAATTTGCTCTGTGTTTCAAAATCCAGAAATGCCAAGAGG	2760
OY	2761	AATGTGCAACATTTCTCTGCCCACTCTGGGTCTTAAAGAAACAAAGTCCAAAGTCACTT	2820
Db	2761	AATGTGCAACATTTCTCTGCCCACTCTGGGTCTTAAAGAAACAAAGTCCAAAGTCACTT	2820
OY	2821	TTGAATGGAACAAAGAGAAAGAAATATCAAGAAAGAAATGAGTCTAATATCAAGCTGTAC	2880
Db	2821	TTGAATGGAACAAAGAGAAAGAAATATCAAGAAAGAAATGAGTCTAATATCAAGCTGTAC	2880
OY	2881	AGACAGTTAATATCACTGCAAGCTTTCCTGTGTGTGTCACAGAAAGATTAAGCCACTTGATA	2940
Db	2881	AGACAGTTAATATCACTGCAAGCTTTCCTGTGTGTGTCACAGAAAGATTAAGCCACTTGATA	2940
OY	2941	ATGCCAAATGTAGTATCAAGAGAGGCTCTAGGTTTGTCTATTCATCTCAGATTCAGAGGCA	3000
Db	2941	ATGCCAAATGTAGTATCAAGAGAGGCTCTAGGTTTGTCTATTCATCTCAGATTCAGAGGCA	3000
OY	3001	ACGAAGCTGGCACTATTAATCCAAATTAACATGAGACTTTTACAAACCCATATGCTATAC	3060
Db	3001	ACGAAGCTGGCACTATTAATCCAAATTAACATGAGACTTTTACAAACCCATATGCTATAC	3060
OY	3061	CACCACTTTTCCCATCAAGTCAATTTGTTAAACTAATGTAAAGAAATCTGCTAGAGG	3120
Db	3061	CACCACTTTTCCCATCAAGTCAATTTGTTAAACTAATGTAAAGAAATCTGCTAGAGG	3120
OY	3121	AAAACTTTGAGGAACATTCAATGTCCACCTGAAGAGAAATGGAATGAGAACATTGCCAA	3180
Db	3121	AAAACTTTGAGGAACATTCAATGTCCACCTGAAGAGAAATGGAATGAGAACATTGCCAA	3180
OY	3181	GTAACAGTGAGACAAATTAAGCCGTATTAACATTAAGAGAAATGTTTAAAGAACCCAGCT	3240
Db	3181	GTAACAGTGAGACAAATTAAGCCGTATTAACATTAAGAGAAATGTTTAAAGAACCCAGCT	3240
OY	3241	CAAGCAATATTAATGAAGTAGTGTCCAGTACTAATGAAGTGGGCTCCAGTATTAATGAAA	3300
Db	3241	CAAGCAATATTAATGAAGTAGTGTCCAGTACTAATGAAGTGGGCTCCAGTATTAATGAAA	3300
OY	3301	TAGGTTCCAGTGATGAAAACATTCACAGCAAGTAGTGAACACAGAGGCGCAAAATTTGA	3360
Db	3301	TAGGTTCCAGTGATGAAAACATTCACAGCAAGTAGTGAACACAGAGGCGCAAAATTTGA	3360

Db	3301	TAGTTCACAGTGAAGAAAACATTACAGCAGAACTAGTAGAAACAGAGGGCCAAAATTGGA	3360
QY	3361	ATGCATCTCTAGATTAGGGGTTTTTGCAACCTGAGGTCCTATAAAACAAGTCTCTGGAA	3420
Db	3361	ATGCTATGCTTAGATTAGGGGTTTTGCAACCTGAGGTCCTATAAAACAAGTCTCTCGGAA	3420
QY	3421	GTAATTGTAGACATCCTCAATAATAAAAAGCAGAATATGAAGAGTACTCAGACTGTTA	3480
Db	3421	GTAATTGTAGACATCCTCAATAATAAAAAGCAGAATATGAAGAGTACTCAGACTGTTA	3480
QY	3481	ATACAGATTTCTCCTCATATCGATTTCAGATTAACCTTGAACAGGCTTATGGGAAGTACTC	3540
Db	3481	ATACAGATTTCTCCTCATATCGATTTCAGATTAACCTTGAACAGGCTTATGGGAAGTACTC	3540
QY	3541	ATGCATCTCAGGTTTTGTTCTGAGACACCTGATGACCTTATAGATATGCTGAATAAAGG	3600
Db	3541	ATGCATCTCAGGTTTTGTTCTGAGACACCTGATGACCTTATAGATATGCTGAATAAAGG	3600
QY	3601	AAGATACTAGTATTGCTGTAATAATGACATTTAAGGAAAGTTCTGCTGTTTTTACCAAAACG	3660
Db	3601	AAGATACTAGTATTGCTGTAATAATGACATTTAAGGAAAGTTCTGCTGTTTTTACCAAAACG	3660
QY	3661	TCCCAAAAGAGAGCTTGAAGAGAGTCTTACCCCTTACCCCTACACATTTGGCTCAGG	3720
Db	3661	TCCCAAAAGAGAGCTTGAAGAGAGTCTTACCCCTTACCCCTACACATTTGGCTCAGG	3720
QY	3721	GTTACCGAAGAGGGGCCAAGAAATTTAGAGTCTCTGAAAGAGAACCTTATCTAGTGAGATG	3780
Db	3721	GTTACCGAAGAGGGGCCAAGAAATTTAGAGTCTCTGAAAGAGAACCTTATCTAGTGAGATG	3780
QY	3781	AAGAGCTTCCCTGCTTCCACACATTTGTTATTTGGTAAAGTAAACAAATATACCTTCTCAGT	3840
Db	3781	AAGAGCTTCCCTGCTTCCACACATTTGTTATTTGGTAAAGTAAACAAATATACCTTCTCAGT	3840
QY	3841	CTACTAGGCATAGCACCCGTTGCTACCGAGTGTCTGTCTAAGAACACAGAGAGAAATTTAT	3900
Db	3841	CTACTAGGCATAGCACCCGTTGCTACCGAGTGTCTGTCTAAGAACACAGAGAGAAATTTAT	3900
QY	3901	TATCTTTTAAAGAAATAGCTTAAATGATGCTGAGTAACACAGGTATATTTGGCAAAAGCATCTC	3960
Db	3901	TATCTTTTAAAGAAATAGCTTAAATGATGCTGAGTAACACAGGTATATTTGGCAAAAGCATCTC	3960
QY	3961	AGGAACATCACCTTAGTGAAGAAACAAATGTCTGCTAGCTTGTCTTCTTACAGTGCA	4020
Db	3961	AGGAACATCACCTTAGTGAAGAAACAAATGTCTGCTAGCTTGTCTTCTTACAGTGCA	4020
QY	4021	GTGAATTTGGAAGACTTGACTCCAAATACAAACACCCAGAGATCTTCTTGTGATTTGTTCTT	4080
Db	4021	GTGAATTTGGAAGACTTGACTCCAAATACAAACACCCAGAGATCTTCTTGTGATTTGTTCTT	4080
QY	4081	CCAAACAATGAGGAGATGATGCTGTAAGACCCAGGGGTTGGTCTGAGTACAGGAATTTGG	4140
Db	4081	CCAAACAATGAGGAGATGATGCTGTAAGACCCAGGGGTTGGTCTGAGTACAGGAATTTGG	4140
QY	4141	TTTTGAGATGATGAAGAAAGAGAGAACGGGCTTGGAAAGAAATTAATCAAGAAAGCAAAAGCA	4200
Db	4141	TTTTGAGATGATGAAGAAAGAGAGAACGGGCTTGGAAAGAAATTAATCAAGAAAGCAAAAGCA	4200
QY	4201	TGGAATTCAACTTAGTGTAAGCAGACATCTGGGTTGAGAGTGAAGAAAGGCTCTCTGAAG	4260
Db	4201	TGGAATTCAACTTAGTGTAAGCAGACATCTGGGTTGAGAGTGAAGAAAGGCTCTCTGAAG	4260
QY	4261	ACTGCTCAGGGCTATCTCTCTCAGAGTGACATTTTAAACCTCAGCAGAGGGGATACCAATGC	4320
Db	4261	ACTGCTCAGGGCTATCTCTCTCAGAGTGACATTTTAAACCTCAGCAGAGGGGATACCAATGC	4320
QY	4321	AACATTAACCTATAAAGCTCCAGAGAGAAATAGGCTGGAACCTAGAGCTGTGTTAAACAGC	4380
Db	4321	AACATTAACCTATAAAGCTCCAGAGAGAAATAGGCTGGAACCTAGAGCTGTGTTAAACAGC	4380
QY	4381	ATGGAGGCACGCTTCTTAACAGCTAACCTTCCATATTAAGTGAAGTCTTCTTGCCTTAGG	4440
Db	4381	ATGGAGGCACGCTTCTTAAACAGCTAACCTTCCATATTAAGTGAAGTCTTCTTGCCTTAGG	4440



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 26, 2003, 16:19:58 ; Search time 7204.45 Seconds  
(without alignments)  
12838.232 Million cell updates/sec

Title: US-09-734-672-3  
Perfect score: 5711  
Sequence: 1 AGCTCCGCTGAGACTCTCTG.....TCCCCACAGACCACTACTGA 5711

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	2061.4	36.1	2101	11 BC012577 Homo sapi
2	2055.6	36.0	2090	11 BC030969 Homo sapi
3	861.4	15.1	962	14 B0068830 AGENCOURT
4	843	14.8	1089	13 BM452288 AGENCOURT
5	811.8	14.2	845	9 AU122476 AU122476
6	761.8	13.3	783	9 AU142729 AU142729

7	740.4	13.0	743	12	BC681276
8	734.6	12.9	899	14 <td>B0422380 AGENCOURT</td>	B0422380 AGENCOURT
9	706.2	12.4	878	14 <td>B0215100 AGENCOURT</td>	B0215100 AGENCOURT
10	688	12.0	702	9 <td>AU125312</td>	AU125312
11	687.6	12.0	747	12 <td>BG777447</td>	BG777447
12	686.6	12.0	739	12 <td>BF508987</td>	BF508987
13	686.4	12.0	987	14 <td>BM800251 AGENCOURT</td>	BM800251 AGENCOURT
14	657	11.5	900	12 <td>BF791668</td>	BF791668
15	651.6	11.4	675	12 <td>BC257190</td>	BC257190
16	639	11.2	675	9 <td>AL704228</td>	AL704228
17	619.2	10.8	921	12 <td>BG178466</td>	BG178466
18	603.8	10.6	630	9 <td>A1992040</td>	A1992040
19	593	10.4	906	14 <td>B0878445</td>	B0878445
20	593	10.4	909	14 <td>B0676829</td>	B0676829
21	593	10.4	933	14 <td>B0683955</td>	B0683955
22	592.6	10.4	800	12 <td>BF794879</td>	BF794879
23	591.4	10.4	627	9 <td>A1589028</td>	A1589028
24	591	10.3	724	9 <td>A1915085</td>	A1915085
25	586	10.3	947	14 <td>B0679749</td>	B0679749
26	581	10.2	910	14 <td>B0677666</td>	B0677666
27	575	10.1	602	10 <td>BE043993</td>	BE043993
28	548.2	9.6	638	13 <td>BM042282</td>	BM042282
29	536.8	9.4	638	10 <td>BE264293</td>	BE264293
30	531.6	9.3	563	10 <td>AW295197</td>	AW295197
31	530.2	9.3	585	10 <td>AW968546</td>	AW968546
32	518	9.1	666	10 <td>AW968720</td>	AW968720
33	499.8	8.8	509	10 <td>AW504244</td>	AW504244
34	492.8	8.6	518	9 <td>AA804632</td>	AA804632
35	491.8	8.6	523	9 <td>AA702344</td>	AA702344
36	479.4	8.6	509	10 <td>BE018878</td>	BE018878
37	479.4	8.4	1220	14 <td>B0214737</td>	B0214737
38	472.2	8.3	487	9 <td>AA812019</td>	AA812019
39	450.6	7.9	575	10 <td>BE564528</td>	BE564528
40	449.8	7.9	453	9 <td>AT684595</td>	AT684595
41	444	7.8	444	9 <td>AA486004</td>	AA486004
42	437.4	7.7	488	9 <td>AA481498</td>	AA481498
43	421.8	7.4	429	9 <td>AA484941</td>	AA484941
44	420.6	7.4	459	9 <td>AT040685</td>	AT040685
45	418.2	7.3	443	14 <td>BQ308670</td>	BQ308670

## ALIGNMENTS

RESULT 1  
LOCUS BC012577 2101 bp mRNA  
DEFINITION Homo sapiens, clone IMAGE:3996658, mRNA.  
ACCESSION BC012577  
VERSION BC012577.1 GI:15214876  
KEYWORDS HTC.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS Strausberg, R.  
TITLE Direct Submission  
JOURNAL Submitted (15-AUG-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 1A03, Bethesda, MD 20892-2590, USA  
REMARK  
COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgaps-femail.nih.gov](mailto:cgaps-femail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: CLONTECH Laboratories, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Institute for Systems Biology  
<http://www.systemsbio.org>  
contact: [amandad@systemsbiology.org](mailto:amandad@systemsbiology.org)  
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia Greene, Mark Keltman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAL Plate: 19 Row: a Column: 18  
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA 91: 6552298  
 This clone has the following problem: incomplete processing.

Location/Qualifiers

source

1. 2101

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_image="399658"

/tissue\_type="Bladder, carcinoma"

/clone\_lib="NIH\_MGC\_53"

/lab\_host="DH10B"

/note="Vector: PDNR-LIB"

BASE COUNT 756 a 385 c 452 g 508 t

ORIGIN

Query Match 36.1%; Score 2061.4; DB 11; Length 2101;

Best Local Similarity 99.4%; Pred. No. 0;

Matches 2079; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

```

1  AGCTGGCTGAGACTTCCTGGACCCGACACGAGCTTGGGGTTCTCAGATACTGGCC 60
6  AGCTGGCTGAGACTTCCTGGACGGGGA-CAGGCTGGGGTTCTCAGATACTGGCC 64
61  CTGGGCTGAGAGGCTTCCACCTCTGCTGGGTAAGTTCATTTGGAACAGAAAGAA 120
65  CCTGGGCTGAGAGGCTTCCACCTCTGCTGGGTAAGTTCATTTGGAACAGAAAGAA 124
121  TGGATTTATCTGCTCTTCCGCTTGAAGAAGTACAAATGTCATTAATGCTATGCA 180
125  TGGATTTATCTGCTCTTCCGCTTGAAGAAGTACAAATGTCATTAATGCTATGCA 184
181  TCTTGAAGTCCCACTGCTGGAGTGTATCAAGAACCTGCTCCAAAGTGTGACC 240
185  TCTTGAAGTCCCACTGCTGGAGTGTATCAAGAACCTGCTCCAAAGTGTGACC 244
241  ACATATTTTGGCAAAATTTTGGCATGCTGAACCTTCTCAACCAAGAAAGAGGCTT 300
245  ACATATTTTGGCAAAATTTTGGCATGCTGAACCTTCTCAACCAAGAAAGAGGCTT 304
301  GTCTTTATGTAAGATGATATACCAAAAGAGCTTACAGAACTACAGATTTAGTC 360
305  GTCTTTATGTAAGATGATATACCAAAAGAGCTTACAGAACTACAGATTTAGTC 364
361  AACTTTGTAAGAGTATTTGAATCATTTGCTTTTCAAGCTGACAGGTTGGAGT 420
365  AACTTTGTAAGAGTATTTGAATCATTTGCTTTTCAAGCTGACAGGTTGGAGT 424
421  ATGCAAGAGCTATATTTTGCAAAAGAGAAATTAATCTCTCTGAACATCTAAAGAT 480
425  ATGCAAGAGCTATATTTTGCAAAAGAGAAATTAATCTCTCTGAACATCTAAAGAT 484
481  AAGTTTATGTAAGATGATATACCAAAAGAGCTTACAGAACTACAGATTTAGTC 540
485  AAGTTTATGTAAGATGATATACCAAAAGAGCTTACAGAACTACAGATTTAGTC 544
541  AACCAGAAATCTTCTGAGAAACAGCTCAGTGTCCAACTCTTAACCTTGGAA 600
545  AACCAGAAATCTTCTGAGAAACAGCTCAGTGTCCAACTCTTAACCTTGGAA 604
601  CTGTGAGAACTCTGAGAAACAGCTCAGTGTCCAACTCTTAACCTTGGAA 660
605  CTGTGAGAACTCTGAGAAACAGCTCAGTGTCCAACTCTTAACCTTGGAA 664
661  AATGGGATCTGATCTTCTGAAGATACCGTTAATAAGCACTTATTCAGTGGAG 720
665  AATGGGATCTGATCTTCTGAAGATACCGTTAATAAGCACTTATTCAGTGGAG 724
721  ATCAAGATCTGATCTTCTGAAGATACCGTTAATAAGCACTTATTCAGTGGAG 780

```

```

725  ATCAAGATCTGATCTTCTGAAGATACCGTTAATAAGCACTTATTCAGTGGAG 784
781  CAAAAGAGCTGCTGATATTTTTCAGAGGATGTAAACAATCTCAACATCTAC 840
785  CAAAAGAGCTGCTGATATTTTTCAGAGGATGTAAACAATCTCAACATCTAC 844
841  CCAGTATATGATTTTGAACACCACTGAGAACGCTGACAGCTGAGAGGATCAGA 900
845  CCAGTATATGATTTTGAACACCACTGAGAACGCTGAGAGGATCAGAAAGT 904
901  ATCAGGATGCTGCTGTTTCAACCTGATGTGAGACCATGTGGCAAAATCTAC 960
905  ATCAGGATGCTGCTGTTTCAACCTGATGTGAGACCATGTGGCAAAATCTAC 964
961  GCTCATATCAGATGAGAACACAGCTTATTTACTACATTAAGACAGAAATGTA 1020
965  GCTCATATCAGATGAGAACACAGCTTATTTACTACATTAAGACAGAAATGTA 1024
1021  AGGCTGAATTTCTGATATTAAGCAACAGCTGCTTGAAGAGACCAACATACAG 1080
1025  AGGCTGAATTTCTGATATTAAGCAACAGCTGCTTGAAGAGACCAACATACAG 1084
1081  GGGCTGGAAGTAAAGAACATGATATGATAGCGGACCTCCAGACAGAAAAAG 1140
1085  GGGCTGGAAGTAAAGAACATGATATGATAGCGGACCTCCAGACAGAAAAAG 1144
1141  ATCTGAATGCTATGCTGCTGCTGAGAGAAAGATGGAATAGAGAACTGGCAT 1200
1145  ATCTGAATGCTATGCTGCTGCTGAGAGAAAGATGGAATAGAGAACTGGCAT 1204
1201  CAGAGATCTGAGATACGAGATGCTTCTTGGATTAACATTAAGAGATTGAGA 1260
1205  CAGAGATCTGAGATACGAGATGCTTCTTGGATTAACATTAAGAGATTGAGA 1264
1261  AAGTTAATGAGGTTTCCAGAGTGAAGTGTGAGTGTGATGATGATGATGAT 1320
1265  AAGTTAATGAGGTTTCCAGAGTGAAGTGTGAGTGTGATGATGATGATGAT 1324
1321  GGGAGTCTGAATCAATGCAAGAGTGTGATGATGATGATGATGATGATGAT 1380
1325  GGGAGTCTGAATCAATGCAAGAGTGTGATGATGATGATGATGATGATGAT 1384
1381  AATATTTCTGCTTCTCAGAGAAATAGACTTACGTGGCAGTGTCTCTAGGCT 1440
1385  AATATTTCTGCTTCTCAGAGAAATAGACTTACGTGGCAGTGTCTCTAGGCT 1444
1441  TATGTAAAGTGAAGAGTGTCTCAGAAATCAGTGAAGTGAAGTGAAGTGA 1500
1445  TATGTAAAGTGAAGAGTGTCTCAGAAATCAGTGAAGTGAAGTGAAGTGA 1504
1501  TTGGGAAAAACCTATCGAAGAGGCAAGCTCCCACTTAAGCATGTAACTGA 1560
1505  TTGGGAAAAACCTATCGAAGAGGCAAGCTCCCACTTAAGCATGTAACTGA 1564
1561  TATTTATAGAGCAATTTGTACTGAGCCAGATATATCAAGAGCTCCCTCA 1620
1565  TATTTATAGAGCAATTTGTACTGAGCCAGATATATCAAGAGCTCCCTCA 1624
1621  AATTAAGGCTTAAAGAGACCTTACATGAGCTTCACTGAGGATTTTATCA 1680
1625  AATTAAGGCTTAAAGAGACCTTACATGAGCTTCACTGAGGATTTTATCA 1684
1681  CAGATTTGGCAGTTCAAAAGACTCTCTGAATGATTAATCAGGGAATCAAC 1740
1685  CAGATTTGGCAGTTCAAAAGACTCTCTGAATGATTAATCAGGGAATCAAC 1744
1741  AGAATGGCAAGTATGATATTAATGATGATGATGATGATGATGATGATGAT 1800
1745  AGAATGGCAAGTATGATATTAATGATGATGATGATGATGATGATGATGAT 1804
1801  CTATTCAGATGAGAAAAATCTTAACCAATGATGATGATGATGATGATGAT 1860
1805  CTATTCAGATGAGAAAAATCTTAACCAATGATGATGATGATGATGATGAT 1864

```

QY 1861 AACGAAAGCTGACCTATTAAGCAGATATAAGCAATATGAACTGCAATTAATATATCC 1920  
 Db 1865 AAACGAAAGCTGACCTATTAAGCAGATATAAGCAATATGAACTGCAATTAATATATCC 1924  
 QY 1921 ACAATTCAAAAGCACCCTAAAGAAATAGGCTGAGAGAGAACTCTTCTACAGCATATTC 1980  
 Db 1925 ACAATTCAAAAGCACCCTAAAGAAATAGGCTGAGAGAGAACTCTTCTACAGCATATTC 1984  
 QY 1981 ATGCGCTGACCTAGTACATAGCAATCTAAGCCCACTAATCTACTGATTCGAAA 2040  
 Db 1985 ATGCGCTGACCTAGTACATAGCAATCTAAGCCCACTAATCTACTGATTCGAAA 2044  
 QY 2041 TTGATAGTGTCTCTAGCAGTGAAGATTAAGAAAAAGTACACCAAA 2091  
 Db 2045 TTGATAGTGTCTCTAGCAGTGAAGATTAAGAAAAAGTACACCAAA 2095  
 RESULT 2  
 BC030969 2090 bp mRNA linear HNC 13-JUN-2002  
 LOCUS Homo sapiens, Similar to breast cancer 1, early onset, clone  
 DEFINITION IMAGE:4804551, mRNA:  
 ACCESSION BC030969  
 VERSION BC030969.1 GI:21411299  
 KEYWORDS HTC.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 2090)  
 AUTHORS Strausberg, R.  
 TITLE Direct Submission.  
 JOURNAL Submitted (03-JUN-2002) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: http://mgc.nci.nih.gov  
 CONTACT: MGC help desk  
 EMAIL: cgabbs-r@mail.nih.gov  
 Tissue Procurement: DCTD/DP  
 CDNA Library Preparation: CLONTECH Laboratories, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
 Center, Stanford University School of Medicine, Stanford, CA 94305  
 Web site: http://www.shgc.stanford.edu  
 CONTACT: (Dickson, Mark) mcd@paxil.stanford.edu  
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.  
 Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov  
 Series: IRL Plate: 41 Row: p Column: 11  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: matched mRNA gi: 6552298  
 This clone has the following problem: frame shifted.  
 FEATURES  
 source  
 1..2090  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4804551"  
 /tissue\_type="prostate, adenocarcinoma."  
 /clone\_lib="NIH\_MGC\_60"  
 /lab\_host="DH10B"  
 /note="vector: pDNR-LIB"  
 BASE COUNT 754 a 385 c 444 g 507 t  
 ORIGIN  
 Query Match 36.0%; Score 2055.6; DB 11; Length 2090;  
 Best Local Similarity 99.3%; Pred. No. 0;  
 Matches 2075; Conservative 0; Mismatches 14; Indels 1; Gaps 1;  
 QY 1 AGCTGCGTAGACCTTCTTGACCCCGCACAGCTGTGGGTTTCTCAGATAACTGGGCC 60

Db 2 AGCTGCGTAGACCTTCTTGACCCCGCACAGCTGTGGGTTTCTCAGATAACTGGGCC 60  
 QY 61 CCTGCGCTGAGAGGCGCTTCACCTCTGCTGCTGGGTAAGTTGATGGAACAGAAAGAA 120  
 Db 61 CCTGCGCTGAGAGGCGCTTCACCTCTGCTGCTGGGTAAGTTGATGGAACAGAAAGAA 120  
 QY 121 TGATATTATCTGCTCTCGGTTGAAGAGTACAAATGTCATTATGCTATGCAGAAA 180  
 Db 121 TGATATTATCTGCTCTCGGTTGAAGAGTACAAATGTCATTATGCTATGCAGAAA 180  
 QY 181 TCTTAGAGTGTCCCATCTGCTGAGTGTATCAAGAAACCTGTCTCACAAACCTGACC 240  
 Db 181 TCTTAGAGTGTCCCATCTGCTGAGTGTATCAAGAAACCTGTCTCACAAACCTGACC 240  
 QY 241 ACATATTCTTGCAAAATTTTGCGATGCTAAACCTGTCAACAGAAAGAAAGGCGCTTCACGT 300  
 Db 241 ACATATTCTTGCAAAATTTTGCGATGCTAAACCTGTCAACAGAAAGAAAGGCGCTTCACGT 300  
 QY 301 GTCCCTTATGTAGAATGATATTAACCAAAAGAGCGCTACAGAAAGTACAGATTAGTC 360  
 Db 301 GTCCCTTATGTAGAATGATATTAACCAAAAGAGCGCTACAGAAAGTACAGATTAGTC 360  
 QY 361 AACCTGTGAAGAGCTATTGAANAATCATTTGCTTTGACGTTGACACAGGTTTGGAGT 420  
 Db 361 AACCTGTGAAGAGCTATTGAANAATCATTTGCTTTGACGTTGACACAGGTTTGGAGT 420  
 QY 421 ATGCAACAGCATATATTTTGGCAAAAAGAAATATACCTCTCTGAAACATCTAAAGATG 480  
 Db 421 ATGCAACAGCATATATTTTGGCAAAAAGAAATATACCTCTCTGAAACATCTAAAGATG 480  
 QY 481 AAGTTCTATCATCCCAAGATGAGGCTACAGAAACCGTGCACAAACACTTCTACAGATG 540  
 Db 481 AAGTTCTATCATCCCAAGATGAGGCTACAGAAACCGTGCACAAACACTTCTACAGATG 540  
 QY 541 AACCCGAAATATCCTTCCTTGGCAGAAACCAAGTCTCAGTCTCAACTCTTAACCTTGGAA 600  
 Db 541 AACCCGAAATATCCTTCCTTGGCAGAAACCAAGTCTCAGTCTCAACTCTTAACCTTGGAA 600  
 QY 601 CTGTGGAACCTGAGAGCAAAAGCAGCGATACACCTCAAAAAGAGCTGTCTTACATTTG 660  
 Db 601 CTGTGGAACCTGAGAGCAAAAGCAGCGATACACCTCAAAAAGAGCTGTCTTACATTTG 660  
 QY 661 AATTGGATCTGATTTCTTGAAGATACCGTTAATAGCAACTTTCAGTGTGGAG 720  
 Db 661 AATTGGATCTGATTTCTTGAAGATACCGTTAATAGCAACTTTCAGTGTGGAG 720  
 QY 721 ATCAGAAATTTGTACAAATCACCCTCAAGGAACCAAGGATGAATCAAGTTGGATTCTG 780  
 Db 721 ATCAGAAATTTGTACAAATCACCCTCAAGGAACCAAGGATGAATCAAGTTGGATTCTG 780  
 QY 781 CAAAAAAGGCTGCTGTAATTTCTGAGACGAGTATCAAAATCTGAAACATCATCAAC 840  
 Db 781 CAAAAAAGGCTGCTGTAATTTCTGAGACGAGTATCAAAATCTGAAACATCATCAAC 840  
 QY 841 CCAGTATATATGATTTGGAACACCACTGAGAGGCTGAGGCGATCCGAAAAGT 900  
 Db 841 CCAGTATATATGATTTGGAACACCACTGAGAGGCTGAGGCGATCCGAAAAGT 900  
 QY 901 ATCAGGATGTTCTGTTCAAACTTGATGAGAGCATGTGGACAAATTAATCATGCA 960  
 Db 901 ATCAGGATGTTCTGTTCAAACTTGATGAGAGCATGTGGACAAATTAATCATGCA 960  
 QY 961 GCTCATTTACAGATGAGAGCAGCTTTATCTACTCTAAAGACAGATGATGTA 1020  
 Db 961 GCTCATTTACAGATGAGAGCAGCTTTATCTACTCTAAAGACAGATGATGTA 1020  
 QY 1021 AGGCTGAATTTCTGTATAAAGCAAAACAGCTGAGCTTACAGAGAGCAACATTAACAGAT 1080  
 Db 1021 AGGCTGAATTTCTGTATAAAGCAAAACAGCTGAGCTTACAGAGAGCAACATTAACAGAT 1080  
 QY 1081 GGGCTGGAAGTAAGAAACATGTAATGATAGCGGACTCCAGCACAGAAAAAAGGTAG 1140  
 Db 1081 GGGCTGGAAGTAAGAAACATGTAATGATAGCGGACTCCAGCACAGAAAAAAGGTAG 1140



```

Db      1081 GGGCTGAGAGTAAGAAACATGTAAATGATAGCGGACCTCCACACAGAAAAAAGTAG 1140
OY      1141 ATCTGAATGCTGATCCCTGTGTGAGAGAAAGATGAATAGCAAACTGCCATGCT 1200
Db      1141 ATCTGAATGCTGATCCCTGTGTGAGAGAAAGATGAATAGCAAACTGCCATGCT 1200
OY      1201 CAGAGAACTCTGAGAGTACTGAGAGATGTTCTTGATTAACACTAAATAGCAGCATTCACA 1260
Db      1201 CAGAGAACTCTGAGAGTACTGAGAGATGTTCTTGATTAACACTAAATAGCAGCATTCACA 1260
OY      1261 AAGTTAAAGTGGTGTCTTCCAGAGTGAACCTGTAGCTTGTGATGACTCAGCATGAG 1320
Db      1261 AAGTTAAAGTGGTGTCTTCCAGAGTGAACCTGTAGCTTGTGATGACTCAGCATGAG 1320
OY      1321 GGGAGCTGCAATCAATGCCAAGTAGCTGATGATTTGAGCGTTCCTAAATAGGTTGAG 1380
Db      1321 GGGAGCTGCAATCAATGCCAAGTAGCTGATGATTTGAGCGTTCCTAAATAGGTTGAG 1380
OY      1381 AATATTTCTGTTCTTCCAGAGAAATAGACTTAAGTCCAGTGTCTCATGAGGCTTTAA 1440
Db      1381 AATATTTCTGTTCTTCCAGAGAAATAGACTTAAGTCCAGTGTCTCATGAGGCTTTAA 1440
OY      1441 TATGTAAAGTGAAGAGTCTCTCAATCAGTGAAGTGAATATTTGAAGACAAATAT 1500
Db      1441 TATGTAAAGTGAAGAGTCTCTCAATCAGTGAAGTGAATATTTGAAGACAAATAT 1500
OY      1501 TTGGGAAACCTATGCGAAGAGGCAAGGCTCCCACTTAAGCCATGTAAGTGAAGAAATC 1560
Db      1501 TTGGGAAACCTATGCGAAGAGGCAAGGCTCCCACTTAAGCCATGTAAGTGAAGAAATC 1560
OY      1561 TAAATATAGAGACATTTGTTACTGAGCCACAGATATATACAAAGGCTCCCTCACAAATA 1620
Db      1561 TAAATATAGAGACATTTGTTACTGAGCCACAGATATATACAAAGGCTCCCTCACAAATA 1620
OY      1621 AATTAAGCCTTAAAGAGAGACTACATCGGCGCTTCATCTGAGGATTTTATACAGAAAG 1680
Db      1621 AATTAAGCCTTAAAGAGAGACTACATCGGCGCTTCATCTGAGGATTTTATACAGAAAG 1680
OY      1681 CAGATTTGGCAGTTCAAAAGAGCTCCTGAATGATTAATCAGGGAACCTAACCAAGGAGC 1740
Db      1681 CAGATTTGGCAGTTCAAAAGAGCTCCTGAATGATTAATCAGGGAACCTAACCAAGGAGC 1740
OY      1741 AGAATGCTCAAGTGAATATTTACTAATAGTGCATAGAGATTAAGCAAAAGGCGATT 1800
Db      1741 AGAATGCTCAAGTGAATATTTACTAATAGTGCATAGAGATTAAGCAAAAGGCGATT 1800
OY      1801 CTATTCAGAAATGAGAAAAATCCTAACCAATAGATTAATCAGGAAAAAGATCGCTTTCA 1860
Db      1801 CTATTCAGAAATGAGAAAAATCCTAACCAATAGATTAATCAGGAAAAAGATCGCTTTCA 1860
OY      1861 AAACGAAAGCTGAACCTATTAAGCAGCAGTATTAAGCAATATGGAACCTCGAATTAATATCC 1920
Db      1861 AAACGAAAGCTGAACCTATTAAGCAGCAGTATTAAGCAATATGGAACCTCGAATTAATATCC 1920
OY      1921 ACAATTCAAAAGACCTTAAAGAAATAGGCTGAGGAGAGTCTTCTACAGGCAATATTC 1980
Db      1921 ACAATTCAAAAGACCTTAAAGAAATAGGCTGAGGAGAGTCTTCTACAGGCAATATTC 1980
OY      1981 ATGGCCTGACTAGTACTAGCAGTAAATCTAAAGCCCACTAATGCTACTGAATTCGAAA 2040
Db      1981 ATGGCCTGACTAGTACTAGCAGTAAATCTAAAGCCCACTAATGCTACTGAATTCGAAA 2040
OY      2041 TTGATAGTTGTTCTAGCAGTGAAGATTAAGAAAAAAGTACAAACA 2090
Db      2041 TTGATAGTTGTTCTAGCAGTGAAGATTAAGAAAAAAGTACAAACA 2090

```

RESULT 3  
 LOCUS B0068830 962 bp mRNA linear EST 02-APR-2002  
 DEFINITION AGENCOURT\_6740238 NIH\_MGC\_47 Homo sapiens cDNA clone IMAGE:5802685  
 ACCESSION B0068830 5' mRNA sequence.

VERSION B0068830.1 GI:19897888  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Mammalia: Euteleostomi: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Primates: Catarrhini: Hominiidae: Homo.  
 REFERENCE 1 (bases 1 to 962)  
 NIH-MGC <http://mhc.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
 Plate: LNCM2037 row: h column: 14  
 High quality sequence stop: 700.

## FEATURES

source Location/Qualifiers  
 1..962

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5802685"  
 /clone\_lib="NIH\_MGC\_47"  
 /tissue\_type="neuroblastoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: brain; Vector: pOTB7; Site\_1: XhoI; Site\_2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(C). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC Library."  
 BASE COUNT 302 a 195 c 235 g 229 t 1 others

Query Match 15.1%; Score 861.4; DB 14; Length 962;  
 Best Local Similarity 96.2%; Pred. No. 1.6e-192;

Matches 890; Conservative 0; Mismatches 33; Indels 2; Gaps 1;

```

OY      4451 TCCAGAACCAACGATCAGAAAAAGCAGTATTAACTTCACAGAAAGTAGTAATCCG 4510
Db      15 TTGGCAGCAGGCGACATCAGAAAAAGCAGTATTAACTTCACAGAAAGTAGTAATCC 74
OY      4511 TATAGCCGAAATCCAGAAAGGCTTTCGCTGACAAAGTTGAGGTCTGCAGATATTC 4570
Db      75 TATAGCCGAAATCCAGAAAGGCTTTCGCTGACAAAGTTGAGGTCTGCAGATATTC 134
OY      4571 TACCAATTAATAAAGAACAGAGTGAAGAGTATCCCTTCTAAATGCCATATY 4630
Db      135 TACCAATTAATAAAGAACAGAGTGAAGAGTATCCCTTCTAAATGCCATATY 194
OY      4631 AGATGATGAGTGTGATGACAGTGTGCTGGGAGTCTTACAGATTAAGAACTACCATC 4690
Db      195 AGATGATGAGTGTGATGACAGTGTGCTGGGAGTCTTACAGATTAAGAACTACCATC 254
OY      4691 TCAAGAGAGCTCATTAAGTGTGATGAGAGGACCAAGCTGGAAGAGTGTGGCC 4750
Db      255 TCAAGAGAGCTCATTAAGTGTGATGAGAGGACCAAGCTGGAAGAGTGTGGCC 314
OY      4751 ACAGCATTTGACGGAACATCTTACTGCCAAGCAAGATCTAAGGAAACCCCTTACT 4810
Db      315 ACAGCATTTGACGGAACATCTTACTGCCAAGCAAGATCTAAGGAAACCCCTTACT 374
OY      4811 GGAATCTGGAATCAGCCTCTCTGATGAGCCGTAATCTGATGAGGACAGC 4870
Db      375 GGAATCTGGAATCAGCCTCTCTGATGAGCCGTAATCTGATGAGGACAGC 434

```

```

QY 4871 CCCAGATCAGCTCGTGTGGCAACAATACCATCTTCACTTCACTTGAAGTTCGCCA 4930
    |||
Db 4435 CCCAGATCAGCTCGTGTGGCAACAATACCATCTTCACTTGAAGTTCGCCA 4934
QY 4931 ATTGAAGTTGCAATCTGCCCAGAGTCCAGTCTGCTCACTACTGATCTCTGG 4990
    |||
Db 495 ATTGAAGTTGCAATCTGCCCAGAGTCCAGTCTGCTCACTACTGATCTCTGG 554
QY 4991 GTATATGCAATGGAAGAAGTGTGAGCAGGGGAGGAAGCAGATTTGACAGCTTCAACAGA 5050
    |||
Db 555 TTATATGCAATGGAAGAAGTGTGAGCAGGGGAGGAAGCAGATTTGACAGCTTCAACAGA 614
QY 5051 AAGGCTCAACAAAGAATGTCATGCTGTGTCTGCTGACCCAGAGAAATTTATGCT 5110
    |||
Db 615 AAGGCTCAACAAAGAATGTCATGCTGTGTCTGCTGACCCAGAGAAATTTATGCT 674
QY 5111 CGGTGCAAGTTGCCAGAAACACCACTTAACTAATCTAATTTCTGAAGAGAC 5170
    |||
Db 675 CGGTGCAAGTTGCCAGAAACACCACTTAACTAATCTAATTTCTGAAGAGAC 734
QY 5171 TACTCATGTTGTATGAAACAGATGCTGAGTTGTGTGAGAGGACACTGAATATTT 5230
    |||
Db 735 TACTCATGTTGTATGAAACAGATGCTGAGTTGTGTGAGAGGACACTGAATATTT 794
QY 5231 TCTAGCAATTTGCGGAGAGAAATGGTAGTATTTCTGAGGAGACCACTGATTTAA 5290
    |||
Db 795 TCTAGCAATTTGCGGAGAGAAATGGTAGTATTTCTGAGGAGACCACTGATTTAA 854
QY 5291 AGAAGCAAAATGCTGAATGAGCATATTTTGAAGTCAGAGAGATGTGG--TCAATGGA 5348
    |||
Db 855 AGAAGCAAAATGCTGAATGAGCATATTTTGAAGTCAGAGAGATGTGGTCAATGGA 914
QY 5349 AGAAGCACCAGGTCACAAAGCGAG 5373
    |||
Db 915 AGAAGCACCAGGTCACAAAGCGAG 939
    |||

RESULT 4
BM452288 1089 bp mRNA linear EST 05-FEB-2002
LOCUS AGENCOURT_6386302 NIH_MGC_72 Homo sapiens cdna clone IMAGE:5356666
DEFINITION 5', mRNA sequence.
ACCESSION BM452288
VERSION BM452288.1 GI:18501328
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
          Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1089)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
          Email: cgabbs-r@mail.nih.gov
          Tissue Procurement: ATCC/DCID/PTP
          CDNA Library Preparation: Life Technologies, Inc.
          DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LNL at:
          http://image.llnl.gov
          Plate: LLM12200 row: k column: 19
          High quality sequence stop: 667.
          Location/Qualifiers
            1..1089
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone="IMAGE:5326666"
              /clone_1lb="NIH_MGC_72"
              /tissue_type="melanotic melanoma"
              /lab_host="DH10B (phage-resistant)"
              /note="Organ: skin; Vector: pCMV-Sport6; Site_1: NotI;

```

```

BASE COUNT 389 a 212 c 248 g 240 t
ORIGIN
Query Match 14.8%; Score 843; DB 13; Length 1089;
Best Local Similarity 96.1%; Pred. No. 3,9e-188;
Matches 929; Conservative 0; Mismatches 30; Indels 8; Gaps 6;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies.
QY 1925 TTCAAAAGCACCTTAAAGAAATAGGCTGAGAGAGAACTTTTACAGCAATATTCATGC 1984
    |||
Db 13 TTCAAAAGCACCTTAAAGAAATAGGCTGAGAGAGAACTTTTACAGCAATATTCATGC 72
QY 1985 GCTTGAACCTAGTACGATGAAATCTAACCCCACTAATTTGATGAAATTTGA 2044
    |||
Db 73 GCTTGAACCTAGTACGATGAAATCTAACCCCACTAATTTGATGAAATTTGA 132
QY 2045 TACTTGTCTAGCAGTGAAGATTAAGAAAGAAAGAAAGTACACCAATGCGCAGCA 2104
    |||
Db 133 TACTTGTCTAGCAGTGAAGATTAAGAAAGAAAGTACACCAATGCGCAGCA 192
QY 2105 CAGCAGAACTTACAACTATGGAAGTAAAGAACTGCACTGAGAGCCAGAAAGCTAA 2164
    |||
Db 193 CAGCAGAACTTACAACTATGGAAGTAAAGAACTGCACTGAGAGCCAGAAAGCTAA 252
QY 2165 CAAGCCAAATGACACAGCAAGTAAAGACATGACATGATCTTTCCAGAGCTGAAGTT 2224
    |||
Db 253 CAAGCCAAATGACACAGCAAGTAAAGACATGACATGATCTTTCCAGAGCTGAAGTT 312
QY 2225 ACAAATGACACCTGCTTCTTCTAAGTGTAAATACAGAGCACTTAAAGATTTGT 2284
    |||
Db 313 ACAAATGACACCTGCTTCTTCTAAGTGTAAATACAGAGCACTTAAAGATTTGT 372
QY 2285 CAATCTAGCTTCCAGAGAGAAAGAAAGAACTGAAAGCTTAAGTGTCTAA 2344
    |||
Db 373 CAATCTAGCTTCCAGAGAGAAAGAAAGAAAGAACTGAAAGCTTAAGTGTCTAA 432
QY 2345 TAATGCTGAAGAGCCCAAGATCTCATGTAAGTGAAGAGGTTTGAAGCTGAAG 2404
    |||
Db 433 TAATGCTGAAGAGCCCAAGATCTCATGTAAGTGAAGAGGTTTGAAGCTGAAG 492
QY 2405 ATCTGTAGAGATGAGATATTTCAATGCTGACTGATGATTAAGGCACTGAGAAAG 2464
    |||
Db 493 ATCTGTAGAGATGAGATATTTCAATGCTGACTGATGATTAAGGCACTGAGAAAG 552
QY 2465 TATCTGTTACTGGAAGTTAGCACTGAGGAGGCAAAACAGAACCAATTAATGTGT 2524
    |||
Db 553 TATCTGTTACTGGAAGTTAGCACTGAGGAGGCAAAACAGAACCAATTAATGTGT 612
QY 2525 GACTCAGTGTGACGATTTGAAAGCCCAAGGAGCAATTCATGCTTCCAAAGTAA 2584
    |||
Db 613 GACTCAGTGTGACGATTTGAAAGCCCAAGGAGCAATTCATGCTTCCAAAGTAA 672
QY 2585 TAAAGATGACACAGAGGCTTTAAGTATCCATTTGGGACATGAGTTAACACAGTGGGA 2644
    |||
Db 673 TAAAGATGACACAGAGGCTTTAAGTATCCATTTGGGACATGAGTTAACACAGTGGGA 732
QY 2645 AACAACATAGAAAT--GGAAGAAAGTGAACCTGATGCTAGATTGTCAGAAATACATCA 2703
    |||
Db 733 AACAACATAGAAATGGAAGAAAGTGAACCTGATGCTAGATTGTCAGAAATACATCA 792
QY 2704 AGTTTCAAGGCGCCAGTCTTTGCTCTGTTTCAAAATCCAGAAATGCAAGAGAAAT 2763
    |||
Db 793 A-GTTTCAAGGCGCCAGTCTTTGCTCTGTTTCAAAATCC-GGAAGTCAAGAGAGAT 850
QY 2764 GTCAACATCTCTGCCACTCTGGTTC--TTAAGAAACAAAGTCCAAAGTCACTTTT 2822
    |||
Db 851 GTCAACATCTCTGGCCACTCTGGGCTTTTAAAGAAACAAAGTCCAAAGTCCCTTTT 910
QY 2823 GAATGTGAACAAA--GGAAGAAATCAAGAAAGATGATCTAA--TATCAAGCCTGT 2878
    |||
Db 911 GAATGTGGGACCAAAAGGACAAAATCCAGGAAAGATGAGATCTAATATTCAGACCTTT 970

```

QY 2879 ACAGACA 2885  
1 1 1 1  
Db 971 AACGAAA 977

RESULT 5  
AUI22476  
LOCUS AUI22476  
DEFINITION AUI22476 MAMMAL Homo sapiens cDNA clone MAMMAL1002447 5', mRNA  
sequence.  
ACCESSION AUI22476  
VERSION AUI22476.1 GI:10937746  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 845)  
Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,  
Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and  
Isogai,T.  
TITLE HRI human cDNA project  
JOURNAL Unpublished (2000)  
COMMENT Contact: Takao Isogai  
Genomics Laboratory  
Helix Research Institute  
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
Tel: 81-438-52-3975  
Fax: 81-438-52-3986  
Email: genomics@hri.co.jp  
HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix  
Research Institute; cDNA library construction; Department of  
Virology, Institute of Medical Science, University of Tokyo, and  
Helix Research Institute.

FEATURES  
source  
1..845  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="MAMMAL1002447"  
/clone\_lib="MAMMAL1"  
/tissue\_type="mammary gland"  
/note="Vector: pME18SFL3"

BASE COUNT 279 a 172 c 191 g 200 t 3 others

ORIGIN

Query Match 14.2% Score 811.8; DB 9; Length 845;  
Best Local Similarity 99.1%; Pred. No. 8.3e-181;  
Matches 836; Conservative 0; Mismatches 5; Indels 3; Gaps 2;

QY 3819 GTAAACAATATACCTTCAGTCTCAGTACGTAGCAGTACCGGTTGCTACCGAGTGTCTCT 3878  
1 GTAAACAATATACCTTCAGTCTCAGTACGTAGCAGTACCGGTTGCTACCGAGTGTCTCT 60.

Db 1

QY 3879 AAGAACAAGAGAGAAATTTATTCATTTGAAGAATAGCTTAAATGACTGCGATACGAG 3938  
1 AAGAACAAGAGAGAAATTTATTCATTTGAAGAATAGCTTAAATGACTGCGATACGAG 120

Db 61

QY 3939 GTAAATATGGCAAGAGCATCTCAGACATCAGTCTAGTGAAGAAACAAATGTTCTGCT 3998  
1 GTAAATATGGCAAGAGCATCTCAGACATCAGTCTAGTGAAGAAACAAATGTTCTGCT 180

Db 121

QY 3999 AGCTGTGTTTCTTCACAGTGAAGTGAAGACTTGCAATATCAACAACCCAG 4058  
1 AGCTGTGTTTCTTCACAGTGAAGTGAAGACTTGCAATATCAACAACCCAG 240

Db 181

QY 4059 GATCCTTTCTTGATTTGGTTCTTCACAAACAATGAGGATCAGTCTGAAGACCGAGAGTT 4118  
1 GATCCTTTCTTGATTTGGTTCTTCACAAACAATGAGGATCAGTCTGAAGACCGAGAGTT 300

Db 241

QY 4119 GGTCGAGTGAACAAGATTTGGTTTCAGATGATGAAGAAAGGAGGCTTGGAAGAA 4178  
1 GGTCGAGTGAACAAGATTTGGTTTCAGATGATGAAGAAAGGAGGCTTGGAAGAA 360

Db 301

QY 4179 AATAATCAAGAGAGCAAAAGCATGTGATTCMAACTTAGTGAAGCAGCATCTGGCTGTGAG 4238  
1 AATAATCAAGAGAGCAAAAGCATGTGATTCMAACTTAGTGAAGCAGCATCTGGCTGTGAG 420

Db 361

QY 4239 AGTGAACAAGCGTCTCTGAGAAGCTGCTGAGGCTATCTCTCAGAGTGACATTTTAAC 4298  
1 AGTGAACAAGCGTCTCTGAGAAGCTGCTGAGGCTATCTCTCAGAGTGACATTTTAAC 480

Db 421

QY 4299 ACTCAGCAGAGGAGATACCATGTGACATTAACCTGATTAAGCTCCAGCAGGAATGCTGAA 4358  
1 ACTCAGCAGAGGAGATACCATGTGACATTAACCTGATTAAGCTCCAGCAGGAATGCTGAA 540

Db 481

QY 4359 CTAGAAGCTGTGTTAGAACAGCATGGAGCCAGCCTTTACAGCTTACCTTCCATGATA 4418  
1 CTAGAAGCTGTGTTAGAACAGCATGGAGCCAGCCTTTACAGCTTACCTTCCATGATA 600

Db 541

QY 4419 AGTGAAGCTGTGTTAGAACAGCATGGAGCCAGCCTTTACAGCTTACCTTCCATGATA 4478  
1 AGTGAAGCTGTGTTAGAACAGCATGGAGCCAGCCTTTACAGCTTACCTTCCATGATA 660

Db 601

QY 4479 GTATTACCTTCACAGAAAAGTGAATACCTATTAAGCCAGATCCAGAGGCTTCT 4538  
1 GTATTACCTTCACAGAAAAGTGAATACCTATTAAGCCAGATCCAGAGGCTTCT 720

Db 661

QY 4539 GCTGACAAAGTTGAGGTGTCTGACAGATAGTTCTACAGTAAATAAATAAGAACACAGCA -G 4596  
1 GCTGACAAAGTTGAGGTGTCTGACAGATAGTTCTACAGTAAATAAATAAGAACACAGCA 780

Db 721

QY 4597 TGGAAAGTCA-TCCCTCTTCTAATGCCATCATTAGATGATGAGTGATGATGATGATGAT 4655  
1 TGGAAAGTCA-TCCCTCTTCTAATGCCATCATTAGATGATGATGATGATGATGATGATGAT 840

Db 781

QY 4656 TGCT 4659  
1 TGCT 844

Db 841

RESULT 6  
AUI42729  
LOCUS AUI42729  
DEFINITION AUI42729 783 bp mRNA linear EST 05-AUG-2002  
sequence.  
ACCESSION AUI42729  
VERSION AUI42729.1 GI:11004250  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 783)  
Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,  
Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and  
Isogai,T.  
TITLE HRI human cDNA project  
JOURNAL Unpublished (2000)  
COMMENT Contact: Takao Isogai  
Genomics Laboratory  
Helix Research Institute  
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
Tel: 81-438-52-3975  
Fax: 81-438-52-3986  
Email: genomics@hri.co.jp  
HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix  
Research Institute; cDNA library construction; Department of  
Virology, Institute of Medical Science, University of Tokyo, and  
Helix Research Institute.

FEATURES  
source  
1..783  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="Y79AA1000792"  
/clone\_lib="Y79AA1"

```

BASE COUNT      255 a      146 c      185 g      194 t      3 others
ORIGIN
Query Match      13.3%; Score 761.8; DB 9; Length 783;
Best Local Similarity 99.2%; Pred. No. 5.4e-169;
Matches 774; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

/cell_type="retinoblastoma"
/cell_line="Y79"
/note="Vector: PME18SFL3"

OY 3579 TTATGATGATGGTGAATTAAGAGATAGTCTTTGCTGGAATAATGACATTAAGAAAGT 3638
Db 1 TTATGATGATGGTGAATTAAGAGATAGTCTTTGCTGGAATAATGACATTAAGAAAGT 60

OY 3639 TCTGCTGTTTTAGCAAAAGCGTCCAGAAAGAGAGAGCTTAGCAGAGCTCTACCCCTTC 3698
Db 61 TCTGCTGTTTTAGCAAAAGCGTCCAGAAAGAGAGAGCTTAGCAGAGCTCTACCCCTTC 120

OY 3699 ACCCATACACATTTGGCTCAGGGTTACCGNAGAGGGGCCAAGAAATTAGAGTCCTAGAA 3758
Db 121 ACCCATACACATTTGGCTCAGGGTTACCGNAGAGGGGCCAAGAAATTAGAGTCCTAGAA 180

OY 3759 GAGACATTTATCTAGTGAAGATGAAGAGCTTCCCTGCTCCACACTTTATTTGGTAAA 3818
Db 181 GAGACATTTATCTAGTGAAGATGAAGAGCTTCCCTGCTCCACACTTTATTTGGTAAA 240

OY 3819 GTAACAAATATACCTTCTCAGTCTAGTACAGGATGACAGCGTTGCTACCGAGTCTGTCT 3878
Db 241 GTAACAAATATACCTTCTCAGTCTAGTACAGGATGACAGCGTTGCTACCGAGTCTGTCT 300

OY 3879 AAGACACAGAGAGAGATTTATATCATTTGAAGAAATAGCTTTAATGACTGCTAGTAACCA 3938
Db 301 AAGACACAGAGAGAGATTTATATCATTTGAAGAAATAGCTTTAATGACTGCTAGTAACCA 360

OY 3939 GTATATTTGGCAAGGATCTCGAGAAATCACCCTTAGTGAGAAACAAATTTCTGTCT 3998
Db 361 GTATATTTGGCAAGGATCTCGAGAAATCACCCTTAGTGAGAAACAAATTTCTGTCT 420

OY 3999 AGCTTTGTTTCTTCCAGTCACTGATGATGGAAGAGCTTGACTGCAATATACCAACCCAG 4058
Db 421 AGCTTTGTTTCTTCCAGTCACTGATGATGGAAGAGCTTGACTGCAATATACCAACCCAG 480

OY 4059 GATCCTTTCTTGAATGTTCTTCCAAACAATGAGGATCAGTCTGAAAGCCAGGAGTT 4118
Db 481 GATCCTTTCTTGAATGTTCTTCCAAACAATGAGGATCAGTCTGAAAGCCAGGAGTT 540

OY 4119 GGCTGAGTGACAGGAATTTGTTTCAATGATGAAGAAAGAGGAGGCTTGGAGAA 4178
Db 541 GGCTGAGTGACAGGAATTTGTTTCAATGATGAAGAAAGAGGAGGCTTGGAGAA 600

OY 4179 AATTAATCAGAAAGAGAAAGCATGATTCAACTTAGTGGAACGACATCTGGGTGAG 4238
Db 601 AATTAATCAGAAAGAGAAAGCATGATTCAACTTAGTGGAACGACATCTGGGTGAG 660

OY 4239 AGTGAACAAGAGCTCTCTGAGAGCTCTCAGGAGTATTCCTCTCAGAGTGACATTTAA-C 4297
Db 661 AGTGAACAAGAGCTCTCTGAGAGCTCTCAGGAGTATTCCTCTCAGAGTGACATTTAAAC 720

OY 4298 CACTGACGAGAGGATACCATGCAACATTAACCTGATTAAGCTCCAGCAGAGAAATGGCTGA 4357
Db 721 CACTGACGAGAGGATACCATGCAACATTAACCTGATTAAGCTCCAGCAGAGAAATGGCTGA 780

RESULT 7
Bg681276 743 bp mRNA linear EST 01-MAY-2001
LOCUS Bg681276 602677125F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4751887 5',
DEFINITION mRNA sequence.
ACCESSION Bg681276
VERSION Bg681276.1 GI:13912673
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

```

```

REFERENCE
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/
TITLE Unpublished (1999)
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: c9apbs-r@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LNA10609 row: a column: 08
high quality sequence stop: 741.

FEATURES
source
1. 743
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NCI_CGAP_Skn4"
/issue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: skin; Vector: pCMV-SPORT6; Site: 1; NotI;
Site: 2; SalI; Cloned unidirectionally. primer: Oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP library."

BASE COUNT      233 a      156 c      182 g      172 t
ORIGIN
Query Match      13.0%; Score 740.4; DB 12; Length 743;
Best Local Similarity 99.9%; Pred. No. 6e-164;
Matches 741; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4724 GGAGCAACAGCTGGAAGAGTCTGGGCCACAGATTGACGGAACAATCTTACTTGCCAAG 4783
Db 1 GGAGCAACAGCTGGAAGAGTCTGGGCCACAGATTGACGGAACAATCTTACTTGCCAAG 60

OY 4784 GCAAGATCTTAGAGAGGAACCCCTTAAGTCTGGAATCTGGAATCAGCTCTCTGATGACC 4843
Db 61 GCAAGATCTTAGAGAGGAACCCCTTAAGTCTGGAATCTGGAATCAGCTCTCTGATGACC 120

OY 4844 TGAATCTGATCCCTTGAAGACAGAGCCCGAGAGTCACTCGTGTGGCAATACATC 4903
Db 121 TGAATCTGATCCCTTGAAGACAGAGCCCGAGAGTCACTCGTGTGGCAATACATC 180

OY 4904 TTCACCTCTGCATTTGAAGTTCCCAATGGAAGTTGAGAAATCTGCCAGAGTCAGG 4963
Db 181 TTCACCTCTGCATTTGAAGTTCCCAATGGAAGTTGAGAAATCTGCCAGAGTCAGG 240

OY 4964 TGGTCTCATACTAGTATCTGCTGGGTATTAATGCAATGAGAAAGTGTGAGCAGGA 5023
Db 241 TGGTCTCATACTAGTATCTGCTGGGTATTAATGCAATGAGAAAGTGTGAGCAGGA 300

OY 5024 GAAGCCAGAAATTAACAGCTTCAACAGAAAGGTCACAAAGAAATGTCATGCTGTGTC 5083
Db 301 GAAGCCAGAAATTAACAGCTTCAACAGAAAGGTCACAAAGAAATGTCATGCTGTGTC 360

OY 5084 TGGCTGTACCCCGAGAGAAATTTATGCTGTGTCAGAGTTTCCAGAAACACACATCAC 5143
Db 361 TGGCTGTACCCCGAGAGAAATTTATGCTGTGTCAGAGTTTCCAGAAACACACATCAC 420

OY 5144 TTTAATTAATCTTAATTAAGAGACTACTGATGTTTATGAAGAAACAGATCTGAGTT 5203
Db 421 TTTAATTAATCTTAATTAAGAGACTACTGATGTTTATGAAGAAACAGATCTGAGTT 480

OY 5204 TGTGTGTAAGGAGACACTGAAATATTTTCTAGAAATTCGGGAGGAAATGGTACTTGG 5263
Db 481 TGTGTGTAAGGAGACACTGAAATATTTTCTAGAAATTCGGGAGGAAATGGTACTTGG 540

OY 5264 CTATTTCTGGGTGACCCAGTCTATTAAGAAAGAAATGCTGAATGACATGATTTGA 5323

```

```

|||||
Db 541 CTAATTTCTGGTGACCCAGCTATATTAAGAAAGAAAAAAGCTGAATGAGCATGATTTTGA 600
|||||
QY 5324 ACTCAGAGAGATGTCGTCATATGAGAAACCCAGAGCTCCAAACGAGCAGAGAAATC 5383
|||||
Db 601 ACTCAGAGAGATGTCGTCATATGAGAAACCCAGAGCTCCAAACGAGCAGAGAAATC 660
|||||
QY 5384 CCAGGACAGAAAGATCTTCAGGGGGCTAGAAATCTGTCATGAGGCCCTTCACCAACAT 5443
|||||
Db 661 CCAGGACAGAAAGATCTTCAGGGGGCTAGAAATCTGTCATGAGGCCCTTCACCAACAT 720
|||||
QY 5444 GCCCAGAGATCACTGGAATGC 5465
|||||
Db 721 GCCCAGAGATCACTGGAATGC 742
|||||

RESULT 8
LOCUS B0422380 899 bp mRNA linear EST 23-MAY-2002
DEFINITION AGENCOURT_7802085 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:6042052
ACCESSION B0422380
VERSION B0422380.1 GI:21117695
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 899)
REFERENCE
AUTHORS NIH-MGC http://mgi.mcl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@b-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM13280 row: n column: 05
High quality sequence stop: 597.
Location/Qualifiers
1..899
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="6042052"
/clone_lib="NIH_MGC_92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: testis; Vector: pCMV-Sport6; Site: 1: NotI;
Site: 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH-MGC library."
BASE COUNT 291 a 182 c 192 g 230 t 4 others
ORIGIN
Query Match 12.9%; Score 734.6; DB 14; Length 899;
Best Local Similarity 97.1%; Pred. No. 1.5e-162;
Matches 824; Conservative 0; Mismatches 11; Indels 14; Gaps 7;

```

```

Db 130 GTACAAATATGCTATTAATGCTATGCAAGAAATCTTAGAGTGTCCATCTGTCTGAGTTG 189
210 ATCAAGCAACCTGTCTCCCAAGAGTGTGACCATATTTTGCATTTTGCATGCTGAA 269
|||||
Db 190 ATCAAGCAACCTGTCTCCCAAGAGTGTGACCATATTTTGCATTTTGCATGCTGAA 249
|||||
QY 270 CTCTCAACCAAGAAAGAGGCTTCACAGTGTCTTTATGTAAATGATATTAACCA 329
|||||
Db 250 CTCTCAACCAAGAAAGAGGCTTCACAGTGTCTTTATGTAAATGATATTAACCA 309
|||||
QY 330 AGAGCCTTCAAGAAAGTACGAGATTAGTCACTTTGTAAGAGCTTTTGAATATCTT 389
|||||
Db 310 AGAGCCTTCAAGAAAGTACGAGATTAGTCACTTTGTAAGAGCTTTTGAATATCTT 369
|||||
QY 390 TGTGCTTTCAGCTTGACACAGAGTTTGGAGTATGCAACAGCTATATTTTGCAGAA 449
|||||
Db 370 TGTGCTTTCAGCTTGACACAGAGTTTGGAGTATGCAACAGCTATATTTTGCAGAA 429
|||||
QY 450 GAAATTAATCTCTCTGACATCTAAAGATGAGTTTCTATCAGCAAGATATGAGCTAC 509
|||||
Db 430 GAAATTAATCTCTCTGACATCTAAAGATGAGTTTCTATCATCAAGATATGAGCTAC 489
|||||
QY 510 AGAAACCTGTCCCAAGAGCTTTACAGAGTGAACCCGAAATCTCTCTGAGGAAAC 569
|||||
Db 490 AGAAACCTGTCCCAAGAGCTTTACAGAGTGAACCCGAAATCTCTCTCTCTCTCT 546
|||||
QY 570 AGTCTAGTGTCCACTCTCTAACCCTTGGAACCTGTGACACCTGTGACCAAGCAGCG 629
|||||
Db 547 AGTCTAGTGTCCACTCTCTAACCCTTGGAACCTGTGACACCTGTGACCAAGCAGCG 606
|||||
QY 630 ATACAACTCTCAAGAGCTGTCTTACATTTGAGATGTGATTTCTTGAAGATAC 689
|||||
Db 607 ATACAACTCTCAAGAGCTGTCTTACATTTGAGATGTGATTTCTTGAAGATAC 666
|||||
QY 690 GTTAATTAAGCACTTATTTGAGAGTGTGAGAGTCAAGATTTGTAATATCAATCCCTCA 749
|||||
Db 667 GTTAATTAAGCACTTATTTGAGAGTGTGAGAGTCAAGATTTGTAATATCAATCCCTCA 726
|||||
QY 750 GGACCA-GGAGTGAATAGTTGATTTGTCGCAAGAAAGCTGC-TTGTGAATTTCTG 807
|||||
Db 727 GGAACCAAGGATGAATATGATTTGATTTGTCGCAAGAAAGCTGC-TTGTGAATTTCTG 786
|||||
QY 808 AGACGATGTAAACA-ATACGTGAATC-ATCAACCCAGTATATGA-TTTGAACACCA 864
|||||
Db 787 AGACGATGTAAACAATATGATTTGATTTGTCGCAAGAAAGCTGC-TTGTGAATTTCTG 846
|||||
QY 865 CTGAGAGC 873
|||||
Db 847 CTGAGAAAC 855
|||||

RESULT 9
LOCUS B0215100 878 bp mRNA linear EST 02-MAY-2002
DEFINITION AGENCOURT_7591049 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6065516
ACCESSION B0215100
VERSION B0215100.1 GI:20396500
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 878)
REFERENCE
AUTHORS NIH-MGC http://mgi.mcl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@b-remail.nih.gov
Tissue Procurement: ATCC/DCFTD/PTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation

```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNt at: <http://image.llnl.gov>  
 Plate: L1AM13341 row: 0 column: 21  
 High quality sequence stop: 669.

## FEATURES

## source

Location/Qualifiers

1. 878

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:6065516"

/clone\_lib="NIH\_MGC\_72"

/tissue\_type="melanotic melanoma"

/lab\_host="DH10B (phage-resistant)"

/note="Organ: skin; Vector: pCMV-Sport6; Site: 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies."

BASE COUNT 281 a 180 c 190 g 227 t  
 ORIGIN

Query Match 12.4% Score 706.2; DB 14; Length 878;  
 Best Local Similarity 98.8%; Pred. No. 7.8e-156;

Matches 754; Conservative 0; Mismatches 3; Indels 6; Gaps 4;

30 CAGGCTGTGGGGTTTCTCAGATTAAGTGGCCCTGCGCTGAGAGGCCCTTACCCTCTGC 89  
 |||||  
 13 CAGGCTGTGGGGTTTCTCAGATTAAGTGGCCCTGCGCTGAGAGGCCCTTACCCTCTGC 72  
 |||||  
 90 TCTGGGTAAGTTTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 149  
 |||||  
 73 TCTGGGTAAGTTTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 132  
 |||||  
 150 GTACAAATGTCAATATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATG 209  
 |||||  
 133 GTACAAATGTCAATATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATG 192  
 |||||  
 210 ATCAGAGACCTGTCTCCACAAAGTGTACACATATTTTGCATTTTGCATGCTGAAA 269  
 |||||  
 193 ATCAGAGACCTGTCTCCACAAAGTGTACACATATTTTGCATTTTGCATGCTGAAA 252  
 |||||  
 270 CTTCCTACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 329  
 |||||  
 253 CTTCCTACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 312  
 |||||  
 330 AGGAGCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 389  
 |||||  
 313 AGGAGCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 372  
 |||||  
 390 TGTGCTTTTCAAGTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 449  
 |||||  
 373 TGTGCTTTTCAAGTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 432  
 |||||  
 450 GAAATTAAGTCTCTGTAACATCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 509  
 |||||  
 433 GAAATTAAGTCTCTGTAACATCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 492  
 |||||  
 510 AGAAGAGCTGCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 569  
 |||||  
 493 AGAAGAGCTGCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 549  
 |||||  
 570 AGTCTCAGTCTCAAGTCTTAACCTTGAAGTGTGAAGTGTGAAGTGTGAAGTGTGAAG 629  
 |||||  
 550 AGTCTCAGTCTCAAGTCTTAACCTTGAAGTGTGAAGTGTGAAGTGTGAAGTGTGAAG 609  
 |||||  
 630 ATACACCTCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 689  
 |||||  
 610 ATACACCTCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 669  
 |||||  
 690 GTTAAATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 748  
 |||||  
 670 GTTAAATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 729  
 |||||  
 749 AGGAACCA-GGAGATGAATCAG-TTTGATTTCTGCAAAAAAG 789

Db 730 AGGAGCAGGAGATGAATCAGTTTGCATCTTCACAAAAAG 772

RESULT 10

AU125312

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (2000)

Contact: Takao Isogai

Genomics Laboratory

Helix Research Institute

1532-3 Yana, Kisaizu, Chiba 292-0812, Japan

Tel: 81-438-52-3975

Fax: 81-438-52-3986

Email: genomicehri.co.jp

HRI human cDNA project: 5' - 3' -end one pass sequencing: Helix

Research Institute; cDNA library construction: Department of

Virology, Institute of Medical Science, University of Tokyo, and

Helix Research Institute.

Location/Qualifiers

1. 702

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="NT2RM4001388"

/clone\_lib="NT2RM4"

/cell\_type="teratocarcinoma"

/cell\_line="NT2"

/note="Vector: pME18SFL3; mRNA from uninduced NT2 neuronal precursor cells"

BASE COUNT 235 a 142 c 163 g 159 t 3 others

ORIGIN

Query Match 12.0% Score 688; DB 9; Length 702;

Best Local Similarity 99.4%; Pred. No. 1.4e-151;

Matches 699; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

3941 AATATGGCAAGGATCTCAGAGATCAACCTTAGTGAAGAGAGAGAGAGAGAGAGAG 4000  
 |||||  
 1 AATATGGCAAGGATCTCAGAGATCAACCTTAGTGAAGAGAGAGAGAGAGAGAGAG 60  
 |||||  
 4001 CTGCTTTTCTTCACAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4060  
 |||||  
 61 CTGCTTTTCTTCACAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120  
 |||||  
 4061 TCCCTTTCTTATTTGGTTCTTCACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4120  
 |||||  
 121 TCCCTTTCTTATTTGGTTCTTCACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180  
 |||||  
 4121 TCTGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4180  
 |||||  
 181 TCTGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240  
 |||||  
 4181 TAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4240  
 |||||  
 241 TAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300  
 |||||  
 4241 TGAACAGAGGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4300







```

Db      1  TTAGCTGACACAGGTTTGGAGTATGCAAAACAGCTAATATTTTGCAAAAAGGAAATAA 60
Oy      458  CTCCTCGACATCTAAAGATGAATTTCTATCATCCAAAGATGGGCTACAGAAACCG 517
Db      61  CTCCTCGACATCTAAAGATGAATTTCTATCATCCAAAGATGGGCTACAGAAACCG 120
Oy      518  TGCCAAAAGACTTCTACAGAGTGAACCCGAAAATCCCTCTCTGCAGAAACAGTCTCAG 577
Db      121  TGCCAAAAGACTTCTACAGAGTGAACCCGAAAATCCCTCTCT--GGAACAGTCTCAG 177
Oy      578  TGTCCAACTCTCTACCTTGGAGAGTGAAGACTCTGAGAGAAAGACCGATACAC 637
Db      178  TGTCCAACTCTCTACCTTGGAGAGTGAAGACTCTGAGAGAAAGACCGATACAC 237
Oy      638  TCAAAAGAGCTGTCTACATTTGAATTTGGATCTGATCTCTGGAAGATACCGTTAATA 697
Db      238  TCAAAAGAGCTGTCTACATTTGAATTTGGATCTGATCTCTGGAAGATACCGTTAATA 697
Oy      698  GGCACCTTATTTGCACTGTGGAGATCAAGAAATTTGTACAAATCACCCCTCAAGAAACAG 757
Db      265  -----
Oy      758  GGATGAATTCAGTTTGATTTCTGCAAAAAGCGCTGCTGTGAATTTTCTGACAGCGATCT 817
Db      265  -----
Oy      818  AACAAATCTGACATCATCAACCCAGTAATATGATTTGAACACCACTGAGAAAGCGTGC 877
Db      295  AACAAATCTGACATCATCAACCCAGTAATATGATTTGAACACCACTGAGAAAGCGTGC 354
Oy      878  AGCTGAGAGGATCCAGAAAAGTATCAGAGTATCTGTTTCAAACTTGATGTGAGGCC 937
Db      355  AGCTGAGAGGATCCAGAAAAGTATCAGAGTATCTGTTTCAAACTTGATGTGAGGCC 414
Oy      938  ATGTGGGCAATATCTCATGCGACGCTCATTCAGATGAGACACAGCTTATCTAC 997
Db      415  ATGTGGGCAATATCTCATGCGACGCTCATTCAGATGAGACACAGCTTATCTAC 474
Oy      998  TAAAGACGAATGAATGTAGAAAAGCGTGAATTTCTGATAAAGCAACAGCGTGC 1057
Db      475  TAAAGACGAATGAATGTAGAAAAGCGTGAATTTCTGATAAAGCAACAGCGTGC 534
Oy      1058  AGCAAGAGCCAACTAAACAGATGGGCTGGAAGTAAGAAACATGATATGAGCGGCAG 1117
Db      535  AGCAAGAGCCAACTAAACAGATGGGCTGGAAGTAAGAAACATGATATGAGCGGCAG 594
Oy      1118  TCCAGACACAGAAAAGAGTAGATCTGATGATCCCTGTGTGAGAGAAAAGATG 1177
Db      595  TCCAGACACAGAAAAGAGTAGATCTGATGATCCCTGTGTGAGAGAAAAGATG 654
Oy      1178  GAATTAAGAGAAAGCTGCCATGCTCAGAGAAATCTGAGATGATGAAGATGCTCTGGAT 1237
Db      655  GAATTAAGAGAAAGCTGCCATGCTCAGAGAAATCTGAGATGATGAAGATGCTCTGGAT 714
Oy      1238  AACACTAAATAGCAGCAATTCAGAAAAGTAAAGTGGTTTCCAGAAAGATGAAGTCT 1297
Db      715  AACACTAAATAGCAGCAATTCAGAAAAGTAAAGTGGTTTCCAGAAAGATGAAGTCT 774
Oy      1298  AGGTTTCTGATACATCAGATGGGAGTCTGAATCAATGCCAAAGTACCT-GATGAT 1356
Db      775  AGGTTTCTGATACATCAGATGGGAGTCTGAATCAATGCCAAAGTACCTGAT 834
Oy      1357  TGACAGTT-CTAAATGAGTAGATG--AATATTCGTGGTCTTCAGAGAAATAGACT 1411
Db      835  TGACAGTTCTTAATGAGTAGATGATATTTCTGGTTCTTTCAGAGAAATAGACT 893

```

```

RESULT 14
BF791668          900 bp  mRNA  linear  EST 12-JAN-2001
LOCUS             60225182.F1 NIH_MGC_84 Homo sapiens cDNA clone IMAGE:4344317 5',
DEFINITION        mRNA sequence.
ACCESSION         BF791668

```

```

VERSION          BF791668.1  GI:12096722
KEYWORDS         EST.
SOURCE           human.
ORGANISM         Homo sapiens
REFERENCE        1 (bases 1 to 900)
AUTHORS          Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE            Eukaryota; Mammalia; Primates; Catarrhini; Homnidae; Homo.
JOURNAL          NIH-MGC http://mgc.nci.nih.gov/
COMMENT          National Institutes of Health, Mammalian Gene Collection (MGC)
                  Unpublished (1999)
                  Contact: Robert Strausberg, Ph.D.
                  Email: cgapbs-r@mail.nih.gov
                  Tissue Procurement: ATCC
                  CDNA Library Preparation: Life Technologies, Inc.
                  CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
                  DNA Sequencing by: Incyte Genomics, Inc.
                  Clone distribution: MGC clone distribution information can be
                  found through the I.M.A.G.E. Consortium/LMNL at:
                  http://image.llnl.gov
                  Plate: LAM9962  row: k  column: 06
                  High quality sequence stop: 678.
FEATURES         Location/Qualifiers
                 1..900
                 /organism="Homo sapiens"
                 /db_xref="taxon:9606"
                 /clone="IMAGE:4344317"
                 /clone_lib="NIH_MGC_84"
                 /tissue_type="adrenal cortex carcinoma, cell line"
                 /lab_host="DH10B (phage-resistant)"
                 /note="Organ: adrenal gland; Vector: pCMV-Sport6; Site:1:
                 NotI; Site:2: SalI; Cloned unidirectionally; oligo-dr
                 primed. Average insert size 1.229 kb. Library enriched for
                 full-length clones and constructed by Life Technologies.
                 Note: this is a NIH_MGC library."
BASE COUNT       283 a 176 c 212 g 229 t
ORIGIN
Query Match
Best Local Similarity 96.5%; Score 657; DB 12; Length 900;
Matches 736; Conservative 0; Mismatches 20; Indels 7; Gaps 6;
Oy      3475  CTGTTAATACAGATTTCTCCATATCTGATTTAGATAGATAGAACAGCTATGAGAA 3534
Db      1  CTGTTAATACAGATTTCTCCATATCTGATTTAGATAGATAGAACAGCTATGAGAA 60
Oy      3535  GTAGTCATGCAATCTCAGGTTTCTGAGACACCTGATGACCTGTAGATGATGTTAAA 3594
Db      61  GTAGTCATGCAATCTCAGGTTTCTGAGACACCTGATGACCTGTAGATGATGTTAAA 120
Oy      3595  TAAAGGAAGATAGTATTTGCTGAAATGACATTAAAGAAAGTTGCTGTTTGA 3654
Db      121  TAAAGGAAGATAGTATTTGCTGAAATGACATTAAAGAAAGTTGCTGTTTGA 180
Oy      3655  AAAGGTCACAGAAAGAGAGCTTAGCAGAGTCTAGCCCTTTACCCATACACATTTGG 3714
Db      181  AAAGGTCACAGAAAGAGAGCTTAGCAGAGTCTAGCCCTTTACCCATACACATTTGG 240
Oy      3715  CTCAGGTTTACCGAAGAGGGCCCAAGAAATTAGAGTCTCAGAGAGAACTTATCTAGTG 3774
Db      241  CTCAGGTTTACCGAAGAGGGCCCAAGAAATTAGAGTCTCAGAGAGAACTTATCTAGTG 300
Oy      3775  AGGATGAAGAGCTTCCCTGCTTCCAACTGTTATTTGTAAGTAAACAATTTACTCT 3834
Db      301  AGGATGAAGAGCTTCCCTGCTTCCAACTGTTATTTGTAAGTAAACAATTTACTCT 359
Oy      3835  CTCAGTCTACTAGGCAATAGACCGCTTGTCTACCGAGTGTCTCTTAAGAACACAGAGAGA 3894
Db      360  CTCAGTCTACTAGGCAATAGACCGCTTGTCTACCGAGTGTCTCTTAAGAACACAGAGAGA 419
Oy      3895  ATTATTTATCATTTGAAGATAGCTTAATGATGATGACAGTACAGAGTAAATGGAAGG 3954
Db      420  ATTATTTATCATTTGAAGATAGCTTAATGATGATGACAGTACAGAGTAAATGGAAGG 478

```

Accession	Sequence	Position
QY	3955 CATCTCGAGGAATCACCCTTGTGTAGGAACAAATGTTCTGACGTTGTTCTTCAC	4014
Db	479 CATCTCGAGGAATCACCCTTGTGTAGGAACAAATGTTCTGACGTTGTTCTTCAC	538
QY	4015 AGTCGACTGAATTGGGAAGACTTGACTGCATAATACAAACACCAGATCTTCTGTATG	4074
Db	539 AGTCGACTGAATTGGGAAGACTTGACTGCATAATACAAACACCAGATCTTCTGTATG	598
QY	4075 GTTCTTCCAAACAATGAGGCATAGTCTGAACCCAGGGAGTTGTCGTGTCATCAAG	4134
Db	599 GTTCTTCCAAACAATGAGGCATAGTCTGAACCCAGG--AGTGTCTGTGATACAGG	656
QY	4135 AATTGTTTCAGATGATGAAGAAAGAGAACGGCTTTGGAAAGAAATATATCAAGAGC	4194
Db	657 AA-TGGTTTCAGATGATGAAGAAAGAGG-ACGGCTTCGACAGAAATATATCAAGAGC	714
QY	4195 AAACATGGATTCAAACTTAGTGAAGACAGCATCTGGTGGA 4237	
Db	715 CAACGATGG-TTCACACTTAGGTAGACAGCTCTGTGTGTGGA 756	

RESULT	15
BG257190	
LOCUS	602337790BP1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4508450 5' ,
DEFINITION	mRNA sequence.
ACCESSION	BG257190
VERSION	BG257190.1 GI:12767019
KEYWORDS	EST.
SOURCE	human.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

1 (bases 1 to 938)  
NIH-MGC <http://mgc.ncl.nih.gov/>  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.

Tissue Procurement: ATCC  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MCC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLM10387 row: b column: 03  
 High quality sequence set: 657.

FEATURES	Location/Qualifiers
source	1. .938

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4508450"
/clone_id="NH_MGC_92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/organ="testis; Vector: pCMVSPORT6; Site:1. NOT1"
/site_2="Sall; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies
Note: this is a NIH_MGC library."

```

BASE COUNT	294 a	193 c	225 g	226 t
ORIGIN				

Query Match	11.48;	Score 651.6;	DB 12;	Length 938;
Best Local Similarity	93.98;	Pred. No. 6.4e-143;		
Matches 712; Conservative	0;	Mismatches 39;	Indels 7;	Gaps 3;

0y 30 CAGGCGTGTGGGTTTCTCAGATAACTGGGCCCCCTGCGCTCAGGAGGCGTTCAACCTTGC 89  
|||||  
10 CAGGCGTGTGGGTTTCTCAGATAACTGGGCCCCCTGCGCTCAGGAGGCGTTCAACCTTGC 69  
Db  
0y 90 TCTGGGTAAAGTTCAATTGGACAGCAAGAAATGGATTATCTGCTCTTCCGGTGTGAAGAA 149

D	b	70	TCGGGTAAGTTCANTGTGSAACAGAAAGAAATGGATTTATCTGCTCTTCGGGTTGAAGAA	129
Q	y	150	GTACAAATATGTCATTAAATGCTATGCAAGAAATCTTAGAGTGTCCATCTGTGTGAGTTG	209
D	b	130	GTACAAAATGTCTAATATGCTATGTCAGAAATCTTAGAGTGTCCCATCTGTGTGAGTTG	189
Q	y	210	ATCAAGGAACCTGTCTCCCAAGAGTGTGACCAATATTTTGCAAAATTTTGATGTCTAAA	265
D	b	190	ATCAAGGAACCTGTCTCCCAAGAGTGTGACCAATATTTTGCAAAATTTTGATGTCTAAA	249
Q	y	270	CTTCTCAACCAAGAAAGAGGCTTCACAGTGTCTTATGTAAAGATGATATACCAA	329
D	b	250	CTTCTCAACCAAGAAAGAGGCTTCACAGTGTCTTATGTAAAGATGATATACCAA	309
Q	y	330	AGGAGCCTACAGAAAGTACGAGATTTAGTCAACTTGTGTAAGACTAATGAAATCATT	389
D	b	310	AGGAGCCTACAGAAAGTACGAGATTTAGTCAACTTGTGTAAGACTAATGAAATCATT	369
Q	y	390	TGTGCTTTTACAGCTTGACACAGGTTTGGAGTATGCAACACACTAATTTTGCAAAAAG	449
D	b	370	TGTGCTTTTACAGCTTGACACAGGTTTGGAGTATGCAACACACTAATTTTGCAAAAAG	429
Q	y	450	GAAATTAACCTCTCCGAAACATCTAAAGATGAGTTTATCATATCCAAAGTATGGCTAC	509
D	b	430	GAAATTAACCTCTCTGAAACATCTAAAGATGAGTTTATCATATCCAAAGTATGGCTAC	489
Q	y	510	AGAAACCGTGCCAAAAGACTTTACAGAGTGAACCCGAAATCTCTCCGTGCGAAACC	565
D	b	490	AGAAACCGTGCCAAAAGACTTTACAGAGTGAACCCGAAATCTCTCCGT	545
Q	y	570	AGCTCAGTGTCCAACTCTCTAACCTTGGAACTGTGAACACTCTGAGACCAACACACGG	629
D	b	547	AGCTCAGTGTCCAACTCTCTAACCTTGGAACTGTGAACACTCTGAGACCAACACACGG	606
Q	y	630	ATCAACACTCCAAAAGACGTGTCTACATGTAATGGAGTCTGATCTTCTG- AAGTATC	688
D	b	607	ATCAACACTCCAAAAGACGTGTCTACATGTAATGGAGTCTGATCTTCTG- AAGTATC	665
Q	y	689	CGTTTAATAGCAACTTATGTGAGTGTGGAGATCAGAATTTGTACAAATCACCCCTCA	748
D	b	664	CGTTTAATAGCAACTTATGTGAGTGTGGAGATCACAAGATGTTCAATATCACCCT	723
Q	y	749	AGGAACCAAGGATGAATACAGTTGGATTGTGCAAAA	786
D	b	724	CAGGAACAGGATGGAATACAGTGTGATCTGCGCAAAA	761

Search completed: June 27, 2003, 04:59:57  
Job time : 7230.45 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 26, 2003, 16:33:51 ; Search time 734.849 Seconds  
(without alignments)  
11536.634 Million cell updates/sec

Title: US-09-734-672-3  
Perfect score: 5711  
Sequence: 1 AGCTCGCTGAGACTTCTCTG.....TCCCCACAGCAGCACTACTGA 5711

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1055720 seqs, 742224136 residues

Total number of hits satisfying chosen parameters: 2111440

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database: Published\_Applications\_NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed.  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	5709.4	100.0	5711	9	US-09-734-672-3
2	5709.4	100.0	5711	9	US-09-982-828-5
3	5704.6	99.9	5711	9	US-09-734-672-5
4	5704.6	99.9	5711	9	US-09-982-828-3
5	5704.6	99.9	5711	9	US-10-022-819-1
6	5703	99.9	5711	9	US-09-734-672-1
7	5703	99.9	5711	9	US-09-982-828-1
8	364.6	6.4	499	9	US-09-911-904-127
9	175.4	3.1	424	10	US-09-864-761-4552
10	147	2.6	147	10	US-09-864-761-21299
11	121	2.1	121	9	US-09-818-875-653
12	121	2.1	121	9	US-09-818-875-654
13	121	2.1	121	9	US-09-818-875-657
14	121	2.1	121	9	US-09-818-875-658
15	121	2.1	121	9	US-09-818-875-661
16	121	2.1	121	9	US-09-818-875-662
17	121	2.1	121	9	US-09-818-875-665
18	121	2.1	121	9	US-09-818-875-666
19	121	2.1	121	9	US-09-818-875-669

20	121	2.1	121	9	US-09-818-875-670	Sequence 670, App
21	121	2.1	121	9 <td>US-09-818-875-673<td>Sequence 673, App</td></td>	US-09-818-875-673 <td>Sequence 673, App</td>	Sequence 673, App
22	121	2.1	121	9 <td>US-09-818-875-674<td>Sequence 674, App</td></td>	US-09-818-875-674 <td>Sequence 674, App</td>	Sequence 674, App
23	121	2.1	121	9 <td>US-09-818-875-677<td>Sequence 677, App</td></td>	US-09-818-875-677 <td>Sequence 677, App</td>	Sequence 677, App
24	121	2.1	121	9 <td>US-09-818-875-678<td>Sequence 678, App</td></td>	US-09-818-875-678 <td>Sequence 678, App</td>	Sequence 678, App
25	121	2.1	121	9 <td>US-09-818-875-681<td>Sequence 681, App</td></td>	US-09-818-875-681 <td>Sequence 681, App</td>	Sequence 681, App
26	121	2.1	121	9 <td>US-09-818-875-682<td>Sequence 682, App</td></td>	US-09-818-875-682 <td>Sequence 682, App</td>	Sequence 682, App
27	121	2.1	121	9 <td>US-09-818-875-685<td>Sequence 685, App</td></td>	US-09-818-875-685 <td>Sequence 685, App</td>	Sequence 685, App
28	121	2.1	121	9 <td>US-09-818-875-686<td>Sequence 686, App</td></td>	US-09-818-875-686 <td>Sequence 686, App</td>	Sequence 686, App
29	121	2.1	121	9 <td>US-09-818-875-689<td>Sequence 689, App</td></td>	US-09-818-875-689 <td>Sequence 689, App</td>	Sequence 689, App
30	121	2.1	121	9 <td>US-09-818-875-690<td>Sequence 690, App</td></td>	US-09-818-875-690 <td>Sequence 690, App</td>	Sequence 690, App
31	121	2.1	121	9 <td>US-09-818-875-693<td>Sequence 693, App</td></td>	US-09-818-875-693 <td>Sequence 693, App</td>	Sequence 693, App
32	121	2.1	121	9 <td>US-09-818-875-694<td>Sequence 694, App</td></td>	US-09-818-875-694 <td>Sequence 694, App</td>	Sequence 694, App
33	121	2.1	121	9 <td>US-09-818-875-697<td>Sequence 697, App</td></td>	US-09-818-875-697 <td>Sequence 697, App</td>	Sequence 697, App
34	121	2.1	121	9 <td>US-09-818-875-698<td>Sequence 698, App</td></td>	US-09-818-875-698 <td>Sequence 698, App</td>	Sequence 698, App
35	121	2.1	121	9 <td>US-09-818-875-701<td>Sequence 701, App</td></td>	US-09-818-875-701 <td>Sequence 701, App</td>	Sequence 701, App
36	121	2.1	121	9 <td>US-09-818-875-702<td>Sequence 702, App</td></td>	US-09-818-875-702 <td>Sequence 702, App</td>	Sequence 702, App
37	121	2.1	121	9 <td>US-09-818-875-705<td>Sequence 705, App</td></td>	US-09-818-875-705 <td>Sequence 705, App</td>	Sequence 705, App
38	121	2.1	121	9 <td>US-09-818-875-706<td>Sequence 706, App</td></td>	US-09-818-875-706 <td>Sequence 706, App</td>	Sequence 706, App
39	121	2.1	121	9 <td>US-09-818-875-709<td>Sequence 709, App</td></td>	US-09-818-875-709 <td>Sequence 709, App</td>	Sequence 709, App
40	121	2.1	121	9 <td>US-09-818-875-710<td>Sequence 710, App</td></td>	US-09-818-875-710 <td>Sequence 710, App</td>	Sequence 710, App
41	121	2.1	121	9 <td>US-09-818-875-713<td>Sequence 713, App</td></td>	US-09-818-875-713 <td>Sequence 713, App</td>	Sequence 713, App
42	121	2.1	121	9 <td>US-09-818-875-714<td>Sequence 714, App</td></td>	US-09-818-875-714 <td>Sequence 714, App</td>	Sequence 714, App
43	121	2.1	121	9 <td>US-09-818-875-717<td>Sequence 717, App</td></td>	US-09-818-875-717 <td>Sequence 717, App</td>	Sequence 717, App
44	121	2.1	121	9 <td>US-09-818-875-718<td>Sequence 718, App</td></td>	US-09-818-875-718 <td>Sequence 718, App</td>	Sequence 718, App
45	121	2.1	121	9 <td>US-09-818-875-721<td>Sequence 721, App</td></td>	US-09-818-875-721 <td>Sequence 721, App</td>	Sequence 721, App

## ALIGNMENTS

RESULT 1  
US-09-734-672-3  
Sequence 3, Application US/09734672  
Publication No. US20020183268A1

GENERAL INFORMATION:  
APPLICANT: MURPHY, Patricia D.  
Allens, Antoinette C.  
Critz, Brenda S.  
Olson, Shert J.  
Schelter, Denise B.  
Zeng, Bin

TITLE OF INVENTION: Coding Sequences of the Human  
BRCAl Gene

NUMBER OF SEQUENCES: 72  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morgan Lewis & Bockius LLP  
STREET: 1111 Pennsylvania Ave., N.W.  
CITY: Washington  
STATE: District of Columbia  
COUNTRY: USA  
ZIP: 20004

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/734, 672  
FILING DATE: 03-Dec-2000

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/966, 436  
FILING DATE: 07 No. US20020183268A1-97  
APPLICATION NUMBER: US 08/598, 591  
FILING DATE: 12-Feb-96

ATTORNEY/AGENT INFORMATION:  
NAME: Michael S. Tuscan  
REGISTRATION NUMBER: 43, 210  
REFERENCE/DOCKET NUMBER: 44921-5055-02-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-739-3000  
TELEFAX: 202-739-3001

```

: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
:   LENGTH: 5711 base pairs
:   TYPE: nucleic acid
:   STRANDEDNESS: No. US20020183268A1 Relevant
:   TOPOLOGY: Linear
: MOLECULE TYPE: cDNA
: ORIGINAL SOURCE:
:   ORGANISM: Homo sapiens
:   STRAIN: BRCA1
:   POSITION IN GENOME:
: CHROMOSOME/SEGMENT: 17
:   MAP POSITION: 17q21
: SEQUENCE DESCRIPTION: SEQ ID NO: 3:
: US-09-734-672-3

Query Match      100.0%; Score 5709.4; DB 9; Length 5711;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5710; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGCTGCTGAGACTTCTCGACCCCGCAGGCTGTGGGGTTTCTCAGATACTGGGC 60
DB 1 AGCTGCTGAGACTTCTCGACCCCGCAGGCTGTGGGGTTTCTCAGATACTGGGC 60
QY 61 CCTGCGCTCAGAGGCGCTTCACTCTGCTGTGGTAAAGTTCATTGGAACAGAAAGAA 120
DB 61 CCTGCGCTCAGAGGCGCTTCACTCTGCTGTGGTAAAGTTCATTGGAACAGAAAGAA 120
QY 121 TGGATTATCTGCTTCGCGGTGGAAGTAACAAATGCTAATATGCTATGAGAAA 180
DB 121 TGGATTATCTGCTTCGCGGTGGAAGTAACAAATGCTAATATGCTATGAGAAA 180
QY 181 TCTTAGAGTGTCCCATCTGTCTGGAAGTTCAGCAAGAACCTGTCTCCCAAGTGTGAC 240
DB 181 TCTTAGAGTGTCCCATCTGTCTGGAAGTTCAGCAAGAACCTGTCTCCCAAGTGTGAC 240
QY 241 ACAATATTTTGCATAATTTTGGATCTCTCAACCTTCAACCAAGAAAGGCGCTTCAAGT 300
DB 241 ACAATATTTTGCATAATTTTGGATCTCTCAACCTTCAACCAAGAAAGGCGCTTCAAGT 300
QY 301 GTCCTTATATGATAGATATATACCAAAAGAGCGCTACAAAGAAAGTACAGATTTAGTC 360
DB 301 GTCCTTATATGATAGATATATACCAAAAGAGCGCTACAAAGAAAGTACAGATTTAGTC 360
QY 361 AACCTTGTGAAGAGCTATTGAAATCATTTTGTCTTTTCACTTGTGACACAGGTTTGGAGT 420
DB 361 AACCTTGTGAAGAGCTATTGAAATCATTTTGTCTTTTCACTTGTGACACAGGTTTGGAGT 420
QY 421 ATGCAAAACAGCTAATTTTGGCAAAAAGGAAATAAATCTCTCTGAACATCTAAGAGATG 480
DB 421 ATGCAAAACAGCTAATTTTGGCAAAAAGGAAATAAATCTCTCTGAACATCTAAGAGATG 480
QY 481 AAGTTCTATCATCCAAAGATATGGGCTACAGAAACCGTCCCAAAAGACTCTACAGAGTG 540
DB 481 AAGTTCTATCATCCAAAGATATGGGCTACAGAAACCGTCCCAAAAGACTCTACAGAGTG 540
QY 541 AACCCGAAAATCCCTTCTTGACAGAAACAGCTCTCAAGTGTCAACTCTCTTAACCTTGGAA 600
DB 541 AACCCGAAAATCCCTTCTTGACAGAAACAGCTCTCAAGTGTCAACTCTCTTAACCTTGGAA 600
QY 601 CTGTGGAAGCTCTGAGAGACAAACAGCGGATACACCTCAAAAGACTCTGTCTACATTTG 660
DB 601 CTGTGGAAGCTCTGAGAGACAAACAGCGGATACACCTCAAAAGACTCTGTCTACATTTG 660
QY 661 AATTGGATCTGATTTCTTCTGAGAGATACGGTAAATAGGCACTTATTTGAGTGGAG 720
DB 661 AATTGGATCTGATTTCTTCTGAGAGATACGGTAAATAGGCACTTATTTGAGTGGAG 720
QY 721 ATCAAGAAATTTTACAAATCACCCCTCAAGAAACAGAGATGAATCAGTTGGATTTCTG 780
DB 721 ATCAAGAAATTTTACAAATCACCCCTCAAGAAACAGAGATGAATCAGTTGGATTTCTG 780
QY 781 CAAAAAAGGCTGCTTGTGAATTTTCTGAGACGAGATTAACAAATCTGAACATCATCAAC 840
DB 781 CAAAAAAGGCTGCTTGTGAATTTTCTGAGACGAGATTAACAAATCTGAACATCATCAAC 840

DB 781 CAAAAAAGGCTGCTTGTGAATTTTCTGAGACGAGATTAACAAATCTGAACATCATCAAC 840
QY 841 CCAATATATATGATTTGAAACACACAGAGGCTCAGCTGAGAGGATCCAGAAAGT 900
DB 841 CCAATATATATGATTTGAAACACACAGAGGCTCAGCTGAGAGGATCCAGAAAGT 900
QY 901 ATCAGGCTAGTCTGTTTCAAACTGATGTGAGCGCATGTGGCAACAAATCTCATGCA 960
DB 901 ATCAGGCTAGTCTGTTTCAAACTGATGTGAGCGCATGTGGCAACAAATCTCATGCA 960
QY 961 GCTCATATACAGCATGAGAACAGAGGTTTATCTACTAATAAGACAGATGAATGTAGAAA 1020
DB 961 GCTCATATACAGCATGAGAACAGAGGTTTATCTACTAATAAGACAGATGAATGTAGAAA 1020
QY 1021 AGGCTGATTTCTGATTAATAAAGCAACACCTGCTTACCAAGAGGCCAATACAT 1080
DB 1021 AGGCTGATTTCTGATTAATAAAGCAACACCTGCTTACCAAGAGGCCAATACAT 1080
QY 1081 GGGCTGGAAGTAAAGAAACATGTAATGATAGCGGACTCCAGCAGAAAAAAGGTAG 1140
DB 1081 GGGCTGGAAGTAAAGAAACATGTAATGATAGCGGACTCCAGCAGAAAAAAGGTAG 1140
QY 1141 ATCTGAATGCTGATCCCTGTGTGAGAGAAAGATGAATGAAGCAAGAACTGCCATGCT 1200
DB 1141 ATCTGAATGCTGATCCCTGTGTGAGAGAAAGATGAATGAAGCAAGAACTGCCATGCT 1200
QY 1201 CAGAGAACTCTAGAGATAGTGAAGATGTCTTGGTAACTAAATACCACTTACGA 1260
DB 1201 CAGAGAACTCTAGAGATAGTGAAGATGTCTTGGTAACTAAATACCACTTACGA 1260
QY 1261 AAGTTAATGAGTGGTTTTCAGAGATGATGAACTGTAGTCTGTGATGCTCAGATGATG 1320
DB 1261 AAGTTAATGAGTGGTTTTCAGAGATGATGAACTGTAGTCTGTGATGCTCAGATGATG 1320
QY 1321 GGGAGTCTGATCAATATGCCAAAGTGTGATGATTTGACGTTCTAATGAGTATG 1380
DB 1321 GGGAGTCTGATCAATATGCCAAAGTGTGATGATTTGACGTTCTAATGAGTATG 1380
QY 1381 AATATCTGCTTCTTACAGAAATATGACTTACTGCGAGTATCTCTCATGAGGCTTTAA 1440
DB 1381 AATATCTGCTTCTTACAGAAATATGACTTACTGCGAGTATCTCTCATGAGGCTTTAA 1440
QY 1441 TATGTAAAGTGAAGAGTTCACCTCAATCACTAGACAGTAAATTTGAAGCAAAATAT 1500
DB 1441 TATGTAAAGTGAAGAGTTCACCTCAATCACTAGACAGTAAATTTGAAGCAAAATAT 1500
QY 1501 TTGGGAAACCTATCGGAAGAAAGCAAGGCTCCCAAACTTAAGCCATGTAAATTC 1560
DB 1501 TTGGGAAACCTATCGGAAGAAAGCAAGGCTCCCAAACTTAAGCCATGTAAATTC 1560
QY 1561 TAATTTATAGAGCATTTTGTACTGAGCCACAGATTAATACAGAGGCTCCCTCACAAATTA 1620
DB 1561 TAATTTATAGAGCATTTTGTACTGAGCCACAGATTAATACAGAGGCTCCCTCACAAATTA 1620
QY 1621 AATTTAAGCGTAAAGAGACCTACATCAGGCTTCATCTCAGAGATTTTATCAAGAAAG 1680
DB 1621 AATTTAAGCGTAAAGAGACCTACATCAGGCTTCATCTCAGAGATTTTATCAAGAAAG 1680
QY 1681 CAGATTTGCGAGTTCAAAAGACTCTGGAATATATATCAGGAACTAACCAAGCGAGC 1740
DB 1681 CAGATTTGCGAGTTCAAAAGACTCTGGAATATATATCAGGAACTAACCAAGCGAGC 1740
QY 1741 AGAATGCTCAAGGATGATATTACTAATAGTGGTATGAGAAATTAACCAAAAGGTT 1800
DB 1741 AGAATGCTCAAGGATGATATTACTAATAGTGGTATGAGAAATTAACCAAAAGGTT 1800
QY 1801 CTATTCAGATGAGAAAAATCTTAACCAATAGATCACTCCAAAAAGATCTGCTTTCA 1860
DB 1801 CTATTCAGATGAGAAAAATCTTAACCAATAGATCACTCCAAAAAGATCTGCTTTCA 1860
QY 1861 AAACGAAGCTGAACTAATAGCAGAGTATAGCAATATGGAACCTGAATTAATATTC 1920
DB 1861 AAACGAAGCTGAACTAATAGCAGAGTATAGCAATATGGAACCTGAATTAATATTC 1920
```

Db 1861 AAGGAAAGCTGAACCTATAACGACAGTATAGCAATATGAACTCGAATTAATATATC 1920  
Qy 1921 ACAATTCAAAAGACACCTAAAAAGATAGGCTGAGAGAGAGTCTTACCGAGCATATTC 1980  
Db 1921 ACAATTCAAAAGACACCTAAAAAGATAGGCTGAGAGAGAGTCTTACCGAGCATATTC 1980  
Qy 1981 ATGGGCTTGAACCTAGTAGTGTAGTAAGATCTAAGCCACCACTAATTTAGTCAATTTGCAAA 2040  
Db 1981 ATGGGCTTGAACCTAGTAGTGTAGTAAGATCTAAGCCACCACTAATTTAGTCAATTTGCAAA 2040  
Qy 2041 TTGATAGTGTGTCTAGAGAGTGAAGATTAAGAAAAAAAGTACACCAATTTGCGAGTCA 2100  
Db 2041 TTGATAGTGTGTCTAGAGAGTGAAGATTAAGAAAAAAAGTACACCAATTTGCGAGTCA 2100  
Qy 2101 GGCACAGAGAAACCTCACTGAGAGGTAAAGAACTGACAGTGGAGCCAAAGAA 2160  
Db 2101 GGCACAGAGAAACCTCACTGAGAGGTAAAGAACTGACAGTGGAGCCAAAGAA 2160  
Qy 2161 GTAAACAAGCCAAATGAACAGACAAAGTAAAGACATGACATGATTAATTTCCAGAGCTGA 2220  
Db 2161 GTAAACAAGCCAAATGAACAGACAAAGTAAAGACATGACATGATTAATTTCCAGAGCTGA 2220  
Qy 2221 AGTTAAACAATGACACCTGCTTCTTTACTAAGTGTCAATATCCAGATGAATTTAAGAT 2280  
Db 2221 AGTTAAACAATGACACCTGCTTCTTTACTAAGTGTCAATATCCAGATGAATTTAAGAT 2280  
Qy 2281 TTGTCATCTAGAGCTTCCAAAGAGAAAGAAAGAACTGAAGAACTTAAAGTGT 2340  
Db 2281 TTGTCATCTAGAGCTTCCAAAGAGAAAGAAAGAAAGAACTGAAGAACTTAAAGTGT 2340  
Qy 2341 CTAATTAATGCTGAAGACCCCAAGATCTCATTTAAGTGAGAAAGGGTTTTCGCAAC 2400  
Db 2341 CTAATTAATGCTGAAGACCCCAAGATCTCATTTAAGTGAGAAAGGGTTTTCGCAAC 2400  
Qy 2401 AAAATCTGTAGAGAGTAGAGCATTTTCATTTGTAAGTGTGTAAGTGTATATAGCACTAG 2460  
Db 2401 AAAATCTGTAGAGAGTAGAGCATTTTCATTTGTAAGTGTGTAAGTGTATATAGCACTAG 2460  
Qy 2461 AAATATCTGTGTCTGTAAGTGTAGCACTCTAGGAAAGGCAAAAACGAACCAATTAAT 2520  
Db 2461 AAATATCTGTGTCTGTAAGTGTAGCACTCTAGGAAAGGCAAAAACGAACCAATTAAT 2520  
Qy 2521 GTGTGAGTCAAGTGTGACAGCATTTTGAACCCCAAGGCACTAATTCATGTTTTCGCAAG 2580  
Db 2521 GTGTGAGTCAAGTGTGACAGCATTTTGAACCCCAAGGCACTAATTCATGTTTTCGCAAG 2580  
Qy 2581 ATATATGAATGACACAGAGGCTTTTAAGTATCCATTGGGACATGAAGTTAACCAAGTGC 2640  
Db 2581 ATATATGAATGACACAGAGGCTTTTAAGTATCCATTGGGACATGAAGTTAACCAAGTGC 2640  
Qy 2641 GGGAAACAGCATGTGAATGGAAGAAAGTGAAGTGTGATGTCAGTATTTGAGAAATCAT 2700  
Db 2641 GGGAAACAGCATGTGAATGGAAGAAAGTGAAGTGTGATGTCAGTATTTGAGAAATCAT 2700  
Qy 2701 TCAAGTGTTCAAAGCGCCAGTCAATTTGCTGTTTTCAAATCCAGAAATGCAAGAGG 2760  
Db 2701 TCAAGTGTTCAAAGCGCCAGTCAATTTGCTGTTTTCAAATCCAGAAATGCAAGAGG 2760  
Qy 2761 AATGTGCAACATTTCTGCGCCACCTCTGGTCTTTAAAGAAACAAGTCCAAAGTCACTT 2820  
Db 2761 AATGTGCAACATTTCTGCGCCACCTCTGGTCTTTAAAGAAACAAGTCCAAAGTCACTT 2820  
Qy 2821 TTGAATGTGAACAAAGAGAAATCAAGAAAGATGATGAATATCAAGCTGTAC 2880  
Db 2821 TTGAATGTGAACAAAGAGAAATCAAGAAAGATGATGAATATCAAGCTGTAC 2880  
Qy 2881 AGACAGTTAATATCACTGACAGGCTTTCCTGTGGTGCAGAAAGATGAAGCAATGTATA 2940  
Db 2881 AGACAGTTAATATCACTGACAGGCTTTCCTGTGGTGCAGAAAGATGAAGCAATGTATA 2940  
Qy 2941 ATGCAAAATGTAGATCAAAAGAGGCTCTAGGTTTGTCTATCATCTCAGTTGAGAGCA 3000  
Db 2941 ATGCAAAATGTAGATCAAAAGAGGCTCTAGGTTTGTCTATCATCTCAGTTGAGAGCA 3000

Qy 3001 ACGAACTGAGCTATTACTCCAAATAAACATGACCTTTTACAAAACCCATATCGTATAC 3060  
Db 3001 ACGAACTGAGCTATTACTCCAAATAAACATGACCTTTTACAAAACCCATATCGTATAC 3060  
Qy 3061 CACCACCTTTTCCATCAAGTCAATTTGTTAAACCTAAATGAAGAAAAATCTGCTAGAGG 3120  
Db 3061 CACCACCTTTTCCATCAAGTCAATTTGTTAAACCTAAATGAAGAAAAATCTGCTAGAGG 3120  
Qy 3121 AAAAATTTGAGAGAACATTCATGTCACCTGGAAGGAAATGGGAATATAGAACTTCCAA 3180  
Db 3121 AAAAATTTGAGAGAACATTCATGTCACCTGGAAGGAAATGGGAATATAGAACTTCCAA 3180  
Qy 3181 GTACAGTGACACAATTAAGCCGTATATTAACATTTAGGAAATGTTTTTAAAGAGCAGCT 3240  
Db 3181 GTACAGTGACACAATTAAGCCGTATATTAACATTTAGGAAATGTTTTTAAAGAGCAGCT 3240  
Qy 3241 CAAGCAATATTAATGAAGTAGGTTCCAGTACTAATGAAGTGGGCTCCAGTATTAATGAA 3300  
Db 3241 CAAGCAATATTAATGAAGTAGGTTCCAGTACTAATGAAGTGGGCTCCAGTATTAATGAA 3300  
Qy 3301 TAGGTTCCAGTATGAAGAAACATTCAGAGCAAGTGAAGAAAGAGAGGCAAAATTTGA 3360  
Db 3301 TAGGTTCCAGTATGAAGAAACATTCAGAGCAAGTGAAGAAAGAGAGGCAAAATTTGA 3360  
Qy 3361 ATGCTATGCTTAGATTTAGGGGTTTTGCAACCTGAGGCTATTAACAAAGTCTCTGGA 3420  
Db 3361 ATGCTATGCTTAGATTTAGGGGTTTTGCAACCTGAGGCTATTAACAAAGTCTCTGGA 3420  
Qy 3421 GTAATTTGAACATCTGAAATTAATAAAGCAAGATATGAAGAGTACTGACACTGTTA 3480  
Db 3421 GTAATTTGAACATCTGAAATTAATAAAGCAAGATATGAAGAGTACTGACACTGTTA 3480  
Qy 3481 ATACAGATTTCTCCATATCTGATTTACAGTAACTTAAGAACACCTATGGAGTATGC 3540  
Db 3481 ATACAGATTTCTCCATATCTGATTTACAGTAACTTAAGAACACCTATGGAGTATGC 3540  
Qy 3541 ATGATCTCAGGTTGTTCTGAGACACCTGATGACCTGTTGATGATGATGATGATGATG 3600  
Db 3541 ATGATCTCAGGTTGTTCTGAGACACCTGATGACCTGTTGATGATGATGATGATGATG 3600  
Qy 3601 AAGATACTAGTTTCTGTAAGAAATGACATTAAGAAAGTCTGCTGTTTTAGCAAAAGCG 3660  
Db 3601 AAGATACTAGTTTCTGTAAGAAATGACATTAAGAAAGTCTGCTGTTTTAGCAAAAGCG 3660  
Qy 3661 TCCAGAAAGAGACCTTAGAGAGTCTTACGCCCTTTCACCATATACATTTGGCTCAGG 3720  
Db 3661 TCCAGAAAGAGACCTTAGAGAGTCTTACGCCCTTTCACCATATACATTTGGCTCAGG 3720  
Qy 3721 GTTACCGAGAGGGGCGCAAGAAATTAAGTGTCTGAGAAAGACCTTATCTGACAGATG 3780  
Db 3721 GTTACCGAGAGGGGCGCAAGAAATTAAGTGTCTGAGAAAGACCTTATCTGACAGATG 3780  
Qy 3781 AAGAGCTTCCCTGCTTCCAAACACTTGTATTTGGTAAAGTAAACATATACCTTCTCAGT 3840  
Db 3781 AAGAGCTTCCCTGCTTCCAAACACTTGTATTTGGTAAAGTAAACATATACCTTCTCAGT 3840  
Qy 3841 CTACTAGGCAATAGCAGCGTGTCTACCGAGTGTCTGTCTAGAACACAGAGAGAAATTTAT 3900  
Db 3841 CTACTAGGCAATAGCAGCGTGTCTACCGAGTGTCTGTCTAGAACACAGAGAGAAATTTAT 3900  
Qy 3901 TATCATTTGAAGAAATGCTTAAATAGTCACTGACATACAGGATATATTTGGCAAAAGCATTC 3960  
Db 3901 TATCATTTGAAGAAATGCTTAAATAGTCACTGACATACAGGATATATTTGGCAAAAGCATTC 3960  
Qy 3961 AGAACAATCACCTTAGTAGAGAAACAAATGTTCTGTAGCTTCTTCCACAGTGA 4020  
Db 3961 AGAACAATCACCTTAGTAGAGAAACAAATGTTCTGTAGCTTCTTCCACAGTGA 4020  
Qy 4021 GTGAATTTGGAAGACTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 4080  
Db 4021 GTGAATTTGGAAGACTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 4080



OY	4081	CCAAACCAATGAGGCATGCTGAAAGCCAGAGAGTTGGTCTGAGTACAAAGAAATGG	4140
Db	4081	CCAAACCAATGAGGCATGCTGAAAGCCAGAGAGTTGGTCTGAGTACAAAGAAATGG	4140
OY	4141	TTTCAGATGATGAAGAAAGAGGAGGAGGCTTGGAGAAATATATCAAGAGCGCAAGCA	4200
Db	4141	TTTCAGATGATGAAGAAAGAGGAGGAGGCTTGGAGAAATATATCAAGAGCGCAAGCA	4200
OY	4201	TGGATTCAAATCTAGGTGAAGCAGCATCTGGGTGTGAAGTGAACCAAGCGTCTTGAAG	4260
Db	4201	TGGATTCAAATCTAGGTGAAGCAGCATCTGGGTGTGAAGTGAACCAAGCGTCTTGAAG	4260
OY	4261	ACTGCTCAGGGCTATCTCTCAGAGTGCATTTTAACCACTCAGCAGAGGGATTACATGC	4320
Db	4261	ACTGCTCAGGGCTATCTCTCAGAGTGCATTTTAACCACTCAGCAGAGGGATTACATGC	4320
OY	4321	AACATTAACCTGATTAAGAGCTCCAGCAGAAATGCGTGAACATGAAGCTGTGTAGAAGC	4380
Db	4321	AACATTAACCTGATTAAGAGCTCCAGCAGAAATGCGTGAACATGAAGCTGTGTAGAAGC	4380
OY	4381	ATGGAGCGCAGCTTCTTAACAGCTACCTTCCATATATAGTGAAGCTTCTGCGCTTGAAG	4440
Db	4381	ATGGAGCGCAGCTTCTTAACAGCTACCTTCCATATATAGTGAAGCTTCTGCGCTTGAAG	4440
OY	4441	ACCTCGAAATTCAGAAACAAAGCAGCATGAGAAAGACAGTATTAATCTTCAGAGAAACTA	4500
Db	4441	ACCTCGAAATTCAGAAACAAAGCAGCATGAGAAAGACAGTATTAATCTTCAGAGAAACTA	4500
OY	4501	GTGAATACCCATATAAGCCAGAAATCCAGAAAGGCTTCTGCTGACAAAGTTTGAGTGTG	4560
Db	4501	GTGAATACCCATATAAGCCAGAAATCCAGAAAGGCTTCTGCTGACAAAGTTTGAGTGTG	4560
OY	4561	CAGATAGTTCTACCGTAAAAATTAAGAACCCAGAGTGGAAAGTCAATCCCTCTTAAT	4620
Db	4561	CAGATAGTTCTACCGTAAAAATTAAGAACCCAGAGTGGAAAGTCAATCCCTCTTAAT	4620
OY	4621	GCCCCATCTTATGATATAGTGGTGTGATGACAGTGTGCTGCGGGAGTGTTCAGAAATAGA	4680
Db	4621	GCCCCATCTTATGATATAGTGGTGTGATGACAGTGTGCTGCGGGAGTGTTCAGAAATAGA	4680
OY	4681	ACTACCCATCTCAAGAGGAGCTCATTTAAGTGTGTGATGTGGAGAGCAACAGCTGGAAG	4740
Db	4681	ACTACCCATCTCAAGAGGAGCTCATTTAAGTGTGTGATGTGGAGAGCAACAGCTGGAAG	4740
OY	4741	AGTCTGGGCGCACGATTTGACGGAACATCTTACTTCCAAAGCAGCATCTAGAGGAA	4800
Db	4741	AGTCTGGGCGCACGATTTGACGGAACATCTTACTTCCAAAGCAGCATCTAGAGGAA	4800
OY	4801	CCCCCTACTGGAATCTGGAATCAGCCCTTCTCTGATGAGCCCTGAATCTATCCTTGTG	4860
Db	4801	CCCCCTACTGGAATCTGGAATCAGCCCTTCTCTGATGAGCCCTGAATCTATCCTTGTG	4860
OY	4861	AAGACAGAGCCCGCAGAGTCAAGTCTGTTGGCAACATACATCTTCAACCTCTGCATTTGA	4920
Db	4861	AAGACAGAGCCCGCAGAGTCAAGTCTGTTGGCAACATACATCTTCAACCTCTGCATTTGA	4920
OY	4921	AAGTTCGCCAATTTGAAGTTCGAGAAATCTGCCCAAGAGTCCAGTCTCTCATACTACTG	4980
Db	4921	AAGTTCGCCAATTTGAAGTTCGAGAAATCTGCCCAAGAGTCCAGTCTCTCATACTACTG	4980
OY	4981	ATACTGCTGGGTATATATGCAATGGAAGAAAGTGTGAGCAGGAGAGCAACAATTTGACAG	5040
Db	4981	ATACTGCTGGGTATATATGCAATGGAAGAAAGTGTGAGCAGGAGAGCAACAATTTGACAG	5040
OY	5041	CTTCAACAGAAAGGCTCAACAAAGAAATGTCCATGAGTGGTGTGCGCTGACCCAGAG	5100
Db	5041	CTTCAACAGAAAGGCTCAACAAAGAAATGTCCATGAGTGGTGTGCGCTGACCCAGAG	5100
OY	5101	AATTTATGCTGTGTACAAAGTTTGGCAGAAAAACACACATCACTTAACTAATCTAATTA	5160
Db	5101	AATTTATGCTGTGTGTACAAAGTTTGGCAGAAAAACACACATCACTTAACTAATCTAATTA	5160
OY	5161	CTGAAGAGACTACTCATCTTCTTATTAAGAAACAGATGCTGAGCTTTGTGTGAACGAGAC	5220

Db	5161	CTGAAGAGCTACTCATGTTGTTATGAAGAACAGATGCTGAGTTTGTTGTGTGAACGACAC	5222
QY	5221	TGAATATTTTCTAGGAATTTGCGGGAGGAAATGGTATGTTAGCTATTTCTGGGTGACCC	5280
Db	5221	TGAAATATTTCTAGGAATTTGCGGGAGGAAATGGTATGTTAGCTATTTCTGGGTGACCC	5280
QY	5281	AGCTCTTTTAAAGAAAGAAAATGCTGMAATGACATGATTTTGAAGTCACAGCAGATGTGG	5340
Db	5281	AGCTCTTTTAAAGAAAGAAAATGCTGMAATGACATGATTTTGAAGTCACAGCAGATGTGG	5340
QY	5341	TCATATGAGAAACCAACCAAGGTCCAAAGCGAGCAGAGAAATCCACGACAGAAAGATCT	5400
Db	5341	TCATATGAGAAACCAACCAAGGTCCAAAGCGAGCAGAGAAATCCACGACAGAAAGATCT	5400
QY	5401	TCAGGGGGCTAGAAAATCTGTGTCTATGTGGCCCTTCACCAACATGCCCACAGATCAATGG	5460
Db	5401	TCAGGGGGCTAGAAAATCTGTGTCTATGTGGCCCTTCACCAACATGCCCACAGATCAATGG	5460
QY	5461	AATGATATGCTACAGCTGTGTGTGCTTCTGTGTGAAGGAGAGCTTCATCATTCACCCCTTG	5520
Db	5461	AATGATATGCTACAGCTGTGTGTGCTTCTGTGTGAAGGAGAGCTTCATCATTCACCCCTTG	5520
QY	5521	GCACAGGTTCACCAATTTGTTGTGTGACAGCAGATGCTGGACAGAGACAAATGCT	5580
Db	5521	GCACAGGTTCACCAATTTGTTGTGTGACAGCAGATGCTGGACAGAGACAAATGCT	5580
QY	5581	TCCATCAATTTGGGAGATGTGTGAGGACACTGTGTGACCCGAGAGTGGGTGTGGACA	5640
Db	5581	TCCATCAATTTGGGAGATGTGTGAGGACACTGTGTGACCCGAGAGTGGGTGTGGACA	5640
QY	5641	GTTATACACTCTACCAAGTCCAGAGCTGGACACTCACTGATATCCCGAGATCCCGACA	5700
Db	5641	GTTATACACTCTACCAAGTCCAGAGCTGGACACTCACTGATATCCCGAGATCCCGACA	5700
QY	5701	GCCACTACTGA 5711	
Db	5701	GCCACTACTGA 5711	

RESULT 2  
 US-09-982-828-5  
 Sequence 5, Application US/09982828  
 Publication NO. US20030022184A1  
 GENERAL INFORMATION:  
 APPLICANT: Murphy, Patricia D.  
 Allen, Antonette C.  
 Alvares, Christopher P.  
 Critz, Brenda S.  
 Olson, Sheri J.  
 Thurber, Denise  
 Zeng, Bin  
 TITLE OF INVENTION: Coding Sequences of the Human  
 BRCA1 Gene  
 NUMBER OF SEQUENCES: 72  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Morgan Lewis & Bockius LLP  
 STREET: 1111 Pennsylvania Avenue N. W.  
 CITY: Washington  
 STATE: DC  
 COUNTRY: USA  
 ZIP: 20004  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/982, 828  
 FILING DATE: 22-Oct-2001  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 09/074,453

FILING DATE: 1998-05-06  
APPLICATION NUMBER: US 08/798,691  
FILING DATE: 1997-02-12  
APPLICATION NUMBER: US 08/598,591  
FILING DATE: 1996-02-12  
ATTORNEY/AGENT INFORMATION:  
NAME: Michael S. Tuscan  
REGISTRATION NUMBER: 43,210  
REFERENCE/DOCKET NUMBER: 44921-5053-01-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-739-3000  
TELEFAX: 202-739-3001  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5711 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: not relevant  
TOPOLOGY: Linear  
MOLECULE TYPE: cDNA  
ORGANISM: Homo sapiens  
STRAIN: BRCAL (cm13)  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: 17  
MAP POSITION: 17q21  
SEQUENCE DESCRIPTION: SEQ ID NO: 5:  
US-09-982-828-5

Query Match 100.0%; Score 5709.4; DB 9; Length 5711;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 5710; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 AGCTGCGTGAGTCTTCTGAGACCCGACCCAGCCGCTGCGGTTTCACAGTACTGGGCC 60  
1 AGCTGCGTGAGTCTTCTGAGACCCGACCCAGCCGCTGCGGTTTCACAGTACTGGGCC 60  
61 CCTGCGCTCAGAGGCGCTTACCCCTGCTGCTGGGTAAAGTTCATTGGAACAGAAAGAA 120  
61 CCTGCGCTCAGAGGCGCTTACCCCTGCTGCTGGGTAAAGTTCATTGGAACAGAAAGAA 120  
121 TGGATTATCTGCTCTTCGCGGTGAAGAAAGTACAAATGTCATTATGCTATCAGAGAA 180  
121 TGGATTATCTGCTCTTCGCGGTGAAGAAAGTACAAATGTCATTATGCTATCAGAGAA 180  
181 TCTAGAGTGCCATCTGCTGAGTGTATCAAGAACTGCTCCACAAGTGTGACC 240  
181 TCTAGAGTGCCATCTGCTGAGTGTATCAAGAACTGCTCCACAAGTGTGACC 240  
181 TCTTAGAGTGCCATCTGCTGAGTGTATCAAGAACTGCTCCACAAGTGTGACC 240  
241 ACATATTTTGCATATTTTGCATGCTGAACTTCTCAACCGAAGAAAGGCGCTTCACAGT 300  
241 ACATATTTTGCATATTTTGCATGCTGAACTTCTCAACCGAAGAAAGGCGCTTCACAGT 300  
301 GTCTTTATGTAAAGATGATATACCAAAAGAGCCTACAGAAAGTACGAGATTTAGTC 360  
301 GTCTTTATGTAAAGATGATATACCAAAAGAGCCTACAGAAAGTACGAGATTTAGTC 360  
361 AACTTGTGAAGAGCTATTGAAATCATTTGCTTTTTCAGTTGACACAGGTTTGGAGT 420  
361 AACTTGTGAAGAGCTATTGAAATCATTTGCTTTTTCAGTTGACACAGGTTTGGAGT 420  
421 ATGCAAAAGAGCTATATTTTGCATATTTTGCATATTTTGCATATTTTGCATATTTTGCAT 480  
421 ATGCAAAAGAGCTATATTTTGCATATTTTGCATATTTTGCATATTTTGCATATTTTGCAT 480  
481 AAGTTTCTATCATCAAAAGATGAGCTACAGAAACCGTGCACAAAGACTTCTACAGAGTG 540  
481 AAGTTTCTATCATCAAAAGATGAGCTACAGAAACCGTGCACAAAGACTTCTACAGAGTG 540  
541 AACCCGAAATCTCTCTTCTGAGAAACCGTGCACAAAGACTTCTACAGAGTG 600  
541 AACCCGAAATCTCTCTTCTGAGAAACCGTGCACAAAGACTTCTACAGAGTG 600  
601 CTGTGAGAACTCTGAGAGCAAAAGAGCTGATACAGCTCAAAAGAGCTGCTGCTACATTG 660

601 CTGTGAGAACTCTGAGAGCAAAAGAGCTGATACAGCTCAAAAGAGCTGCTGCTACATTG 660  
661 AATTGGAGTGTGATTTCTTCTGAGAGTATACCGTTATATAGCACTTATTTGAGTGGAG 720  
661 AATTGGAGTGTGATTTCTTCTGAGAGTATACCGTTATATAGCACTTATTTGAGTGGAG 720  
721 ATACAGATTTGTTACCAATATACCCCTCAAGAACCGAGGATGAAATCATGTTGATTTCTG 780  
721 ATACAGATTTGTTACCAATATACCCCTCAAGAACCGAGGATGAAATCATGTTGATTTCTG 780  
781 CAAAAAGGCTGCTTGTGAATTTTCTGAGAGGATGTAACAAATACTGAACATCATCAAC 840  
781 CAAAAAGGCTGCTTGTGAATTTTCTGAGAGGATGTAACAAATACTGAACATCATCAAC 840  
841 CCAGTATATATGATTTGAAACACCACTGAGAGGCTGAGAGGATGTAACAAATACTGAAC 900  
841 CCAGTATATATGATTTGAAACACCACTGAGAGGCTGAGAGGATGTAACAAATACTGAAC 900  
901 ATCAGGATGATTTGTTTCAAACTGATGAGAGGATGTAACAAATACTGAACATCATCAAC 960  
901 ATCAGGATGATTTGTTTCAAACTGATGAGAGGATGTAACAAATACTGAACATCATCAAC 960  
961 GCTCATTTACGATGAGAGCAAGATTTATTTACTCACTAAAGACAGATGATGTAAGAA 1020  
961 GCTCATTTACGATGAGAGCAAGATTTATTTACTCACTAAAGACAGATGATGTAAGAA 1020  
1021 AGGCTGATTTCTGATTAATAAAGCAAGAGCTGCTGATGAGAGGATGTAACAAATACTGA 1080  
1021 AGGCTGATTTCTGATTAATAAAGCAAGAGCTGCTGATGAGAGGATGTAACAAATACTGA 1080  
1081 GGGCTGGAAGTGAAGAAACATGTAATGATGAGGAGATCCACACAGAAAGAAAGTAG 1140  
1081 GGGCTGGAAGTGAAGAAACATGTAATGATGAGGAGATCCACACAGAAAGAAAGTAG 1140  
1141 ATCTGAATGCTGATCCCTGCTGAGAGAAAGATGGAATGAAGCAAACTGCAATGCT 1200  
1141 ATCTGAATGCTGATCCCTGCTGAGAGAAAGATGGAATGAAGCAAACTGCAATGCT 1200  
1201 CAGAGAACTCTAAGAGTACTGAGAGATGTTCTTGGATACACTTAATAGCAGATTTAGA 1260  
1201 CAGAGAACTCTAAGAGTACTGAGAGATGTTCTTGGATACACTTAATAGCAGATTTAGA 1260  
1261 AAGTTATGATGATGTTTCCAGAGAGTATGATGATGATGATGATGATGATGATGATGATG 1320  
1261 AAGTTATGATGATGTTTCCAGAGAGTATGATGATGATGATGATGATGATGATGATGATG 1320  
1321 GGGAGTCTGATCAAAATGCCAAAGTATGATGATGATGATGATGATGATGATGATGATG 1380  
1321 GGGAGTCTGATCAAAATGCCAAAGTATGATGATGATGATGATGATGATGATGATGATG 1380  
1381 AATATTCTGATTTCTTCCAGAGAAATGATGATGATGATGATGATGATGATGATGATGATG 1440  
1381 AATATTCTGATTTCTTCCAGAGAAATGATGATGATGATGATGATGATGATGATGATGATG 1440  
1441 TATGTAAAGTGAAGAGTCTCACTCCAAATCAGTAGAGATATATATGAGAGCAAAATAT 1500  
1441 TATGTAAAGTGAAGAGTCTCACTCCAAATCAGTAGAGATATATATGAGAGCAAAATAT 1500  
1501 TTGGGAAACCTATCGGAAAGGCAAGCCTCCCACTTAAGCATGATGATGATGATGATG 1560  
1501 TTGGGAAACCTATCGGAAAGGCAAGCCTCCCACTTAAGCATGATGATGATGATGATGATG 1560  
1561 TAATTTATGAGAGATTTGTTTCTGAGAGGATGATGATGATGATGATGATGATGATGATG 1620  
1561 TAATTTATGAGAGATTTGTTTCTGAGAGGATGATGATGATGATGATGATGATGATGATG 1620  
1621 AATTTAAAGCTTAAGAGAGCTACATCAGAGGCTTCTGATGAGATTTTATGAGAGAG 1680  
1621 AATTTAAAGCTTAAGAGAGCTACATCAGAGGCTTCTGATGAGATTTTATGAGAGAG 1680  
1681 CAGATTTGAGAGTCAAAAGAGCTGTAATGATGATGATGATGATGATGATGATGATGATG 1740  
1681 CAGATTTGAGAGTCAAAAGAGCTGTAATGATGATGATGATGATGATGATGATGATGATG 1740

Dp	1681	CAGATTTTGGCACTTCCAAAAGACTCTCGAATATGTAATATCAGGAACTACCAACGGAGC	1740
Qy	1741	ACAATGCTCAAGTGAATGAATATTACTAAATAGTGTGATGAGATTAACAAAGGTGATT	1800
Dp	1741	AGAAATGGTCAAGTGTGATGAATATTACTAAATAGTGTGATGAGATTAACAAAGGTGATT	1800
Qy	1801	CAATTACAGATAGAAAAAATCTTAACCAATGATTCACCTGAAAAAACAATCTGCTTCA	1860
Dp	1801	CAATTACAGATAGAAAAAATCTTAACCAATGATTCACCTGAAAAAACAATCTGCTTCA	1860
Qy	1861	AAAGCAAAAGCTGAACCTTAATAGCAGCAGTATAGCAATATGGAATCTGAATTAATATCC	1920
Dp	1861	AAAGCAAAAGCTGAACCTTAATAGCAGCAGTATAGCAATATGGAATCTGAATTAATATCC	1920
Qy	1921	ACAATTTCAAAACCACTTAATAAAGATAGCTGAGAGAAAGAACTCTCCACGAGATATTC	1980
Dp	1921	ACAATTTCAAAACCACTTAATAAAGATAGCTGAGAGAAAGAACTCTCCACGAGATATTC	1980
Qy	1981	ATGCGCTTGAACCTAGTACTAGTAAATAATCTAAGCCCACTTAATTTGTACTAAATGGCAA	2040
Dp	1981	ATGCGCTTGAACCTAGTACTAGTAAATAATCTAAGCCCACTTAATTTGTACTAAATGGCAA	2040
Qy	2041	TTGATATGTTGTTCTAGCAGTGAAGATTAAGAAAAAAATACACCAATGGCAGTCA	2100
Dp	2041	TTGATATGTTGTTCTAGCAGTGAAGATTAAGAAAAAAATACACCAATGGCAGTCA	2100
Qy	2101	GGCAGACAGAAACCTTACACATCAATGAGAAGTAAAGAACTGCAACTGGACCCAAAGAGA	2160
Dp	2101	GGCAGACAGAAACCTTACACATCAATGAGAAGTAAAGAACTGCAACTGGACCCAAAGAGA	2160
Qy	2161	GTAACAAACCAATATGACAGACAGTAAGAATGACATGATATCTTCCACAGCTGA	2220
Dp	2161	GTAACAAACCAATATGACAGACAGTAAGAATGACATGATATCTTCCACAGCTGA	2220
Qy	2221	AGTTAACAAATGACACCTGGTCTCTTTACTAAGTGTCAAAATACCAAGTAACTTAAGAT	2280
Dp	2221	AGTTAACAAATGACACCTGGTCTCTTTACTAAGTGTCAAAATACCAAGTAACTTAAGAT	2280
Qy	2281	TTGTCAATCTTAGCCTTCCAAAGAGAAAAAGAGAGAAACTGAAGAACTTAAGTGT	2340
Dp	2281	TTGTCAATCTTAGCCTTCCAAAGAGAAAAAGAGAGAAACTGAAGAACTTAAGTGT	2340
Qy	2341	CTAATTAATGCTGAAACACCCCAAAGTCTCAATGTTAAGTGGAGAAAGGTTTGGAAACTG	2400
Dp	2341	CTAATTAATGCTGAAACACCCCAAAGTCTCAATGTTAAGTGGAGAAAGGTTTGGAAACTG	2400
Qy	2401	AAAGATCTGTAGAGAGTAGAGTAATTTCAATGGTACCTGTAATGATTTATGGCACTAGG	2460
Dp	2401	AAAGATCTGTAGAGAGTAGAGTAATTTCAATGGTACCTGTAATGATTTATGGCACTAGG	2460
Qy	2461	AAAGATCTGCTTACTGGAAGTGTAGCACTCTAGGGAAGCAAAAAACGAACCAATTAAT	2520
Dp	2461	AAAGATCTGCTTACTGGAAGTGTAGCACTCTAGGGAAGCAAAAAACGAACCAATTAAT	2520
Qy	2521	GTGTGAGTCAGTGTGACAGCAATTTGAAAAAACCCTAAGGAGCAATCATGGTGTCCAAAG	2580
Dp	2521	GTGTGAGTCAGTGTGACAGCAATTTGAAAAAACCCTAAGGAGCAATCATGGTGTCCAAAG	2580
Qy	2581	ATAATAGAAATGACACAGAGAGGCTTTAAATATTCATTGGGACATGAAGTTTACCACAGTC	2640
Dp	2581	ATAATAGAAATGACACAGAGAGGCTTTAAATATTCATTGGGACATGAAGTTTACCACAGTC	2640
Qy	2641	GGGAACCAAGATAGAAATGGAAGAAAGTGAATGTAATGCTACAGTATTTGACAGATCAT	2700
Dp	2641	GGGAACCAAGATAGAAATGGAAGAAAGTGAATGTAATGCTACAGTATTTGACAGATCAT	2700
Qy	2701	TCAAGGTTTCAAAAGCCGACAGTCATTTGCTGTGTTTCAAAATCCAGGAATGACGAAGAGG	2760
Dp	2701	TCAAGGTTTCAAAAGCCGACAGTCATTTGCTGTGTTTCAAAATCCAGGAATGACGAAGAGG	2760
Qy	2761	AATGTGCAACATTTCTGTGCGCAGCTCTGGGTCTTTAAAGAAACAAGTCCAAAGTACTCT	2820
Dp	2761	AATGTGCAACATTTCTGTGCGCAGCTCTGGGTCTTTAAAGAAACAAGTCCAAAGTACTCT	2820

OY	2821	TTGAAATGGAACAAAGSAGAAATACAGSAAAGAAATGAGCTAATATCAAGCTTAC	2880
Db	2821	TTGAAATGGAACAAAGSAGAAATACAGSAAAGAAATGAGCTAATATCAAGCTTAC	2880
OY	2881	AGACAGTTAATATACACAGCGCTTCCGTGCTGTCAGAAAGATAGCCAGTTGATA	2940
Db	2881	AGACAGTTAATATACACAGCGCTTCCGTGCTGTCAGAAAGATAGCCAGTTGATA	2940
OY	2941	ATGCCAAATAGTATATCAAAAGAGGCTCTAGGTTTTGTCTATCATCTAGTTCAGAGCA	3000
Db	2941	ATGCCAAATAGTATATCAAAAGAGGCTCTAGGTTTTGTCTATCATCTAGTTCAGAGCA	3000
OY	3001	ACGAAATGTGACTCATTTACCTCCAAATTAACATGAGACTTTTCAAAAACCATATGCTATAC	3060
Db	3001	ACGAAATGTGACTCATTTACCTCCAAATTAACATGAGACTTTTCAAAAACCATATGCTATAC	3060
OY	3061	CACCACCTTTTCCCATCAAGTCATTTGTTAAACTAAATGTATAGAAAATCTGCTAGAGG	3120
Db	3061	CACCACCTTTTCCCATCAAGTCATTTGTTAAACTAAATGTATAGAAAATCTGCTAGAGG	3120
OY	3121	AAACCTTTGGGAAACCTTCATATGTCACCGGAAGGAAGAAATGGGAAATGAGAACCTTCCAA	3180
Db	3121	AAACCTTTGGGAAACCTTCATATGTCACCGGAAGGAAGAAATGGGAAATGAGAACCTTCCAA	3180
OY	3181	GTACAGTGAACACAAATTAGCCGTAATATACATTAGAGAAAATGTTTTTAAAGAACCCAGCT	3240
Db	3181	GTACAGTGAACACAAATTAGCCGTAATATACATTAGAGAAAATGTTTTTAAAGAACCCAGCT	3240
OY	3241	CAGCAATATTAATGAAGTAGTTCACAGTACTATATGAAGTGGCTCCACGATTTATATATA	3300
Db	3241	CAGCAATATTAATGAAGTAGTTCACAGTACTATATGAAGTGGCTCCACGATTTATATATA	3300
OY	3301	TAGGTTCCAGTGTATGAAGCAATTCATCAAGCAGACTAGTAGAAGACAGAGGCCCAAAATTTGA	3360
Db	3301	TAGGTTCCAGTGTATGAAGCAATTCATCAAGCAGACTAGTAGAAGACAGAGGCCCAAAATTTGA	3360
OY	3361	ATGCTATGCTTAGATTAGGGGTTTTTGCAACCTGAGGCTATTAACAAAGTCTTCGGA	3420
Db	3361	ATGCTATGCTTAGATTAGGGGTTTTTGCAACCTGAGGCTATTAACAAAGTCTTCGGA	3420
OY	3421	GTAATTTAAGCAATCCGGAATTAAGAAACCAAGATATGAAGAAAGTGGTTCAGCTTTA	3480
Db	3421	GTAATTTAAGCAATCCGGAATTAAGAAACCAAGATATGAAGAAAGTGGTTCAGCTTTA	3480
OY	3481	ATACAGATTTCTCTGCATATCTGATTTAGATTAAGAACAGCCATATGGGAAGTAGTC	3540
Db	3481	ATACAGATTTCTCTGCATATCTGATTTAGATTAAGAACAGCCATATGGGAAGTAGTC	3540
OY	3541	ATGCAATCTCAGGTTTCTCTCGAGACACCGATGACCTGTATGATATGTCGAATTAAGG	3600
Db	3541	ATGCAATCTCAGGTTTCTCTCGAGACACCGATGACCTGTATGATATGTCGAATTAAGG	3600
OY	3601	AAGATATCTAGTWTTCGTGAAGATGACATTAAAGAAAGTTCGTGTTTTTAAGCAAAAGCG	3660
Db	3601	AAGATATCTAGTWTTCGTGAAGATGACATTAAAGAAAGTTCGTGTTTTTAAGCAAAAGCG	3660
OY	3661	TCCAGAAAGSAGAGCTTAGCAGAGAGCTCCAGCCCTTCCACCATACACATTTGGCTCAGG	3720
Db	3661	TCCAGAAAGSAGAGCTTAGCAGAGAGCTCCAGCCCTTCCACCATACACATTTGGCTCAGG	3720
OY	3721	GTTAACGGAAGAGGGCCCAAGAAATTAAGATGCTCAGGAAGAACTTATCTAGTAGAGATG	3780
Db	3721	GTTAACGGAAGAGGGCCCAAGAAATTAAGATGCTCAGGAAGAACTTATCTAGTAGAGATG	3780
OY	3781	AAGAGCTTCCCTGCTCCCAACACTGTTATTTGGTAAAGTAAACAAATATCTTCACAGT	3840
Db	3781	AAGAGCTTCCCTGCTCCCAACACTGTTATTTGGTAAAGTAAACAAATATCTTCACAGT	3840
OY	3841	CTACTATAGCAATAGCACCTGTTGCTACCGAGTGTCTCTTACAGACACAGAGAGAAATTTAT	3900
Db	3841	CTACTATAGCAATAGCACCTGTTGCTACCGAGTGTCTCTTACAGACACAGAGAGAAATTTAT	3900

QY 3901 TATCATGAGATAGTCTTAATGACCTGACCTAACAAGATATATGCGCAAGGCAATCTC 3960  
 Db 3901 TATCATGAGATAGTCTTAATGACCTGACCTAACAAGATATATGCGCAAGGCAATCTC 3960  
 QY 3961 AGGAATCATCCTAGTGGAGAAACAAATGTTCTGTAGTGTCTTCTTCTCAGAGCA 4020  
 Db 3961 AGGAATCATCCTAGTGGAGAAACAAATGTTCTGTAGTGTCTTCTTCTCAGAGCA 4020  
 QY 4021 GTGATTTGGAAGACTTGTGCAATATCAAAACACCCAGATCCCTTCTGTTGTTCTT 4080  
 Db 4021 GTGATTTGGAAGACTTGTGCAATATCAAAACACCCAGATCCCTTCTGTTGTTCTT 4080  
 QY 4081 CCAATCAATGAGGACATCTGTAAGACGAGGAGTTGTTCTGAGTACAGGAATGG 4140  
 Db 4081 CCAATCAATGAGGACATCTGTAAGACGAGGAGTTGTTCTGAGTACAGGAATGG 4140  
 QY 4141 TTTGATGATGAAAGAAAGGAGGAGGCTTGGAAAGAAATATCAAGAAAGAAAGCA 4200  
 Db 4141 TTTGATGATGAAAGAAAGGAGGAGGCTTGGAAAGAAATATCAAGAAAGAAAGCA 4200  
 QY 4201 TGGATTTCAACTTGTGTAAGACGACATCTGGGTGTGAGTGAAGCAAGGCTCTGAG 4260  
 Db 4201 TGGATTTCAACTTGTGTAAGACGACATCTGGGTGTGAGTGAAGCAAGGCTCTGAG 4260  
 QY 4261 ACTGCTCAGGCTATCTCTCAGAGTACATTTTAACTTCAAGCAGAGGAGTACATGC 4320  
 Db 4261 ACTGCTCAGGCTATCTCTCAGAGTACATTTTAACTTCAAGCAGAGGAGTACATGC 4320  
 QY 4321 AACATTAACCTGATTAACCTCCAGAGAAATGCTGAACCTTGAAGCTGTGTAAACAGC 4380  
 Db 4321 AACATTAACCTGATTAACCTCCAGAGAAATGCTGAACCTTGAAGCTGTGTAAACAGC 4380  
 QY 4381 ATGGAGACGACCTTCTTAACAGTACCTTCCATCAATGAGTCTTCTGCTGAGG 4440  
 Db 4381 ATGGAGACGACCTTCTTAACAGTACCTTCCATCAATGAGTCTTCTGCTGAGG 4440  
 QY 4441 ACCTGCAAAATCCAGAAACAAAGCAGATCAGAAAAAGCAATTAACCTCAGCAAGAA 4500  
 Db 4441 ACCTGCAAAATCCAGAAACAAAGCAGATCAGAAAAAGCAATTAACCTCAGCAAGAA 4500  
 QY 4501 GTGATTAACCTTATTAAGCAGAAATCCAGAGGCTTCTGCTGACAAATTTGAGTGTG 4560  
 Db 4501 GTGATTAACCTTATTAAGCAGAAATCCAGAGGCTTCTGCTGACAAATTTGAGTGTG 4560  
 QY 4561 CAGATAGTCTTACAGTAAATTAAGAACAGAGAGTGAAGGTCATCCCTCTTAAT 4620  
 Db 4561 CAGATAGTCTTACAGTAAATTAAGAACAGAGAGTGAAGGTCATCCCTCTTAAT 4620  
 QY 4621 GCCCATATTAGATGATGAGTGTGATGACAGCTTCTGAGGAGTCTTCAAGATGAA 4680  
 Db 4621 GCCCATATTAGATGATGAGTGTGATGACAGCTTCTGAGGAGTCTTCAAGATGAA 4680  
 QY 4681 ACTTACCATTCTCAAGAGAGCTCTTAAAGTGTGATGTGAGAGACCAACAGCTGGAAG 4740  
 Db 4681 ACTTACCATTCTCAAGAGAGCTCTTAAAGTGTGATGTGAGAGACCAACAGCTGGAAG 4740  
 QY 4741 AGTGTGGGACACAGATTTGACGGAACATCTTACTTCCCAAGCAGAGATCTAGAGGAA 4800  
 Db 4741 AGTGTGGGACACAGATTTGACGGAACATCTTACTTCCCAAGCAGAGATCTAGAGGAA 4800  
 QY 4801 CCCCTTACCTGGAATCTGGAATCAGCTCTTCTGATGAGCCTTGAATCTGATCTTCTG 4860  
 Db 4801 CCCCTTACCTGGAATCTGGAATCAGCTCTTCTGATGAGCCTTGAATCTGATCTTCTG 4860  
 QY 4861 AAGACAGAGCCAGAGTCTGCTGTGTGCAACATACATCTTCAACCTCTGCAATGGA 4920  
 Db 4861 AAGACAGAGCCAGAGTCTGCTGTGTGCAACATACATCTTCAACCTCTGCAATGGA 4920  
 QY 4921 AAGTTCCTCAATGAAAGTGTGCAAGATCTCCAGAGTCCAGCTGCTCATCTACTG 4980  
 Db 4921 AAGTTCCTCAATGAAAGTGTGCAAGATCTCCAGAGTCCAGCTGCTCATCTACTG 4980  
 QY 4981 ATACTGCTGGTATTAATGCAATGGAAGAGTGTGAGCAGGAGGAAAGCCAGAAATGACAG 5040

Db 4981 ATACTGCTGGTATTAATGCAATGGAAGAGTGTGAGCAGGAGGAGCAAGCAATTTGACAG 5040  
 QY 5041 CTTCAACGAAAGGCTCAACAAAGAAATGCTCATGCTGTGCTGAGCCCTGAGCCAGAG 5100  
 Db 5041 CTTCAACGAAAGGCTCAACAAAGAAATGCTCATGCTGTGCTGAGCCCTGAGCCAGAG 5100  
 QY 5101 AATTATGCTGCTGATCAAGTTTGCAGAAAAACACCATCACTTAACTTAATCTTA 5160  
 Db 5101 AATTATGCTGCTGATCAAGTTTGCAGAAAAACACCATCACTTAACTTAATCTTA 5160  
 QY 5161 CTGAGAGACTTACTGATGTTTGTATGAACAGATGCTGAGTTTGTGTGTAACGAGAC 5220  
 Db 5161 CTGAGAGACTTACTGATGTTTGTATGAACAGATGCTGAGTTTGTGTGTAACGAGAC 5220  
 QY 5221 TGAATATTTTCTGGAATTCGAGGAGGAAATGAGTGTGATGATTTCTGAGGAGACC 5280  
 Db 5221 TGAATATTTTCTGGAATTCGAGGAGGAAATGAGTGTGATGATTTCTGAGGAGACC 5280  
 QY 5281 AGCTATTTAAAGAAAGAAATGCTGATGAGCATGATTTGAAAGTCAAGAGAGATG 5340  
 Db 5281 AGCTATTTAAAGAAAGAAATGCTGATGAGCATGATTTGAAAGTCAAGAGAGATG 5340  
 QY 5341 TCAATGGAAGAACCCAGAGTCCAAAGCAGAGAGAGATTTCCAGAGACAGAAAGATCT 5400  
 Db 5341 TCAATGGAAGAACCCAGAGTCCAAAGCAGAGAGAGATTTCCAGAGACAGAAAGATCT 5400  
 QY 5401 TCAGAGGCTGGAATGCTGATGAGGCTTCTGATGAGGCTTCCACATGCTCCAGATCACTG 5460  
 Db 5401 TCAGAGGCTGGAATGCTGATGAGGCTTCTGATGAGGCTTCCACATGCTCCAGATCACTG 5460  
 QY 5461 AATGATGCTGACAGCTGTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5520  
 Db 5461 AATGATGCTGACAGCTGTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5520  
 QY 5521 GCACAGGTGTCACCCCAATTTGTTGTGTGACAGAGATGCTGGAAGAGAGAGAGTGT 5580  
 Db 5521 GCACAGGTGTCACCCCAATTTGTTGTGTGACAGAGATGCTGGAAGAGAGAGAGTGT 5580  
 QY 5581 TCCATGCAATTTGGGACAGATGCTGAGGACCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 5640  
 Db 5581 TCCATGCAATTTGGGACAGATGCTGAGGACCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 5640  
 QY 5641 GTGTAGCAGCTTACAGAGTCCAGAGAGTCCAGAGTCCAGAGTCCAGAGTCCAGAGT 5700  
 Db 5641 GTGTAGCAGCTTACAGAGTCCAGAGAGTCCAGAGTCCAGAGTCCAGAGTCCAGAGT 5700  
 QY 5701 GCCACTACTGA 5711  
 Db 5701 GCCACTACTGA 5711

RESULT 3  
 US-09-734-672-5  
 ; Sequence 5, Application US/09734672  
 ; Publication No. US20020183268A1  
 GENERAL INFORMATION:  
 APPLICANT: Murphy, Patricia D.  
 Allen, Antonette C.  
 Alvares, Christopher P.  
 Critz, Brenda S.  
 Olson, Sheri J.  
 Schelter, Denise B.  
 Zeng, Bin  
 TITLE OF INVENTION: Coding Sequences of the Human  
 BRCA1 Gene  
 NUMBER OF SEQUENCES: 72  
 CORRESPONDENCE ADDRESSES:  
 ADDRESSEE: Morgan Lewis & Bockius LLP  
 STREET: 1111 Pennsylvania Ave., N.W.  
 CITY: Washington  
 STATE: District of Columbia  
 COUNTRY: USA

ZIP: 20004  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/734,672  
 FILING DATE: 03-Dec-2000  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/966,436  
 FILING DATE: 07-No. US20020183268A1-97  
 APPLICATION NUMBER: US 08/598,591  
 FILING DATE: 12-Feb-96  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Michael S. Tuscan  
 REGISTRATION NUMBER: 43,210  
 REFERENCE/DOCKET NUMBER: 44921-5055-02-US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-739-3000  
 TELEFAX: 202-739-3001  
 INFORMATION FOR SEQ ID NO: 5:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 5711 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: No. US20020183268A1 Relevant  
 TOPOLOGY: Linear  
 MOLECULE TYPE: cDNA  
 ORIGINAL SOURCE:  
 ORGANISM: Homo sapiens  
 STRAIN: BRCA1  
 POSITION IN GENOME:  
 CHROMOSOME/SEGMENT: 17  
 MAP POSITION: 17q21  
 SEQUENCE DESCRIPTION: SEQ ID NO: 5:  
 us-09-734-672-5

Query Match 99.9%; Score 5704.6; DB 9; Length 5711;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 5707; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AGCTGCTAGAGACTTCTGAGCCCGCAGCAGCTGTGGGTTTCTCAGTAATCGGGCC 60  
 DB 1 AACTGCTAGAGACTTCTGAGCCCGCAGCAGCTGTGGGTTTCTCAGTAATCGGGCC 60  
 QY 61 CCTGCGCTAGAGAGGCTTCAACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120  
 DB 61 CCTGCGCTAGAGAGGCTTCAACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120  
 QY 121 TGGATTATCTGCTCTTGGCGTTGAGAGATACAAATGCTAATGCTATGCTATGCTAGAAA 180  
 DB 121 TGGATTATCTGCTCTTGGCGTTGAGAGATACAAATGCTAATGCTATGCTATGCTAGAAA 180  
 QY 181 TCTTAGAGTGTCCATCTGCTGAGAGTGTATGATCAAGAACCTGCTCCACAAAGTGTGACC 240  
 DB 181 TCTTAGAGTGTCCATCTGCTGAGAGTGTATGATCAAGAACCTGCTCCACAAAGTGTGACC 240  
 QY 241 ACATATTTTGAATTTTGCATGCTGAACCTTCTCAACCGAAGAAAGGCGCTTCAAGT 300  
 DB 241 ACATATTTTGAATTTTGCATGCTGAACCTTCTCAACCGAAGAAAGGCGCTTCAAGT 300  
 QY 301 GTCTTTATGTAGAAATGATATTAACCAAGAGAGCTTACAAAGATAGAGATTTAGTC 360  
 DB 301 GTCTTTATGTAGAAATGATATTAACCAAGAGAGCTTACAAAGATAGAGATTTAGTC 360  
 QY 361 AACTGTTGAGAGAGCTATGAGAAATATTTGCTTTTCAAGCTTGAACAGGTTTGGAGT 420  
 DB 361 AACTGTTGAGAGAGCTATGAGAAATATTTGCTTTTCAAGCTTGAACAGGTTTGGAGT 420  
 QY 421 ATGCAAAAGCTATATTTTGAAGAAAGAAATTAAGTCTCTGAAACATCTAAAGATG 480  
 DB 421 ATGCAAAAGCTATATTTTGAAGAAAGAAATTAAGTCTCTGAAACATCTAAAGATG 480

QY 481 AAGTTCTATCATCAAGATATGGCTAGAGAACCGTCCCAAGAGCTTCTACAGAGTG 540  
 DB 481 AAGTTCTATCATCAAGATATGGCTAGAGAACCGTCCCAAGAGCTTCTACAGAGTG 540  
 QY 541 AACCCGAAATCTCTCTGCGAGAAACAGTCTAGTGTCCAACTCTTAACCTTTGAA 600  
 DB 541 AACCCGAAATCTCTCTGCGAGAAACAGTCTAGTGTCCAACTCTTAACCTTTGAA 600  
 QY 601 CTGTGAGAACTGTGAGAGCAACAGCGGATACAACTGAAAGAGCGTGTCTCAATTG 660  
 DB 601 CTGTGAGAACTGTGAGAGCAACAGCGGATACAACTGAAAGAGCGTGTCTCAATTG 660  
 QY 661 AATTGGATCTGATTTCTTGAAGATACCGTTAATTAAGCAACTTAATTCAGTGTGGAG 720  
 DB 661 AATTGGATCTGATTTCTTGAAGATACCGTTAATTAAGCAACTTAATTCAGTGTGGAG 720  
 QY 721 ATCAGAAATGTTCACAAATCACCCTCAAGAAACAGGATGAATCACTTTGATCTG 780  
 DB 721 ATCAGAAATGTTCACAAATCACCCTCAAGAAACAGGATGAATCACTTTGATCTG 780  
 QY 781 CAAAAAAGCGCTGTGTAATTTTCTGAGAGGATGTAACAAATCTGACATCTCAAC 840  
 DB 781 CAAAAAAGCGCTGTGTAATTTTCTGAGAGGATGTAACAAATCTGACATCTCAAC 840  
 QY 841 CCAGTAATATGATTTGAACACCACTAGAGAGCGTGCAGTGAAGAGCATCCAGAAAGT 900  
 DB 841 CCAGTAATATGATTTGAACACCACTAGAGAGCGTGCAGTGAAGAGCATCCAGAAAGT 900  
 QY 901 ATCAGGATGTTCTGTTTCAAACTTGCATGTGAGAGCATGTGGCACAATATCTCATGCCA 960  
 DB 901 ATCAGGATGTTCTGTTTCAAACTTGCATGTGAGAGCATGTGGCACAATATCTCATGCCA 960  
 QY 961 GCTCATTAACGATGAGAGAGAGAGTATTTCTGAGAGGATGTAACAAATCTGACATCTCAAC 1020  
 DB 961 GCTCATTAACGATGAGAGAGAGAGTATTTCTGAGAGGATGTAACAAATCTGACATCTCAAC 1020  
 QY 1021 AGGCTGAATCTGTAAATTAAGCAACAGCTGTGCTTACAGAGAGCCAACTACAGAT 1080  
 DB 1021 AGGCTGAATCTGTAAATTAAGCAACAGCTGTGCTTACAGAGAGCCAACTACAGAT 1080  
 QY 1081 GGGCTGGAATGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140  
 DB 1081 GGGCTGGAATGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140  
 QY 1141 ATCTGAATGCTGATCCCTGTGAGAGAAAGATGAATGAAGCAAGATGATGATGATG 1200  
 DB 1141 ATCTGAATGCTGATCCCTGTGAGAGAAAGATGAATGAAGCAAGATGATGATGATG 1200  
 QY 1201 CAGAGAACTCTAGAGATGATGAGATGATGATGATGATGATGATGATGATGATGATG 1260  
 DB 1201 CAGAGAACTCTAGAGATGATGAGATGATGATGATGATGATGATGATGATGATGATG 1260  
 QY 1261 AAGTTAATGAGTGTGTTTCCAGAGATGATGAATGTTAGTGTGATGATGATGATGATG 1320  
 DB 1261 AAGTTAATGAGTGTGTTTCCAGAGATGATGAATGTTAGTGTGATGATGATGATGATG 1320  
 QY 1321 GGGAGTCTGAATTAATGCAAGAGTGTGATGATGATGATGATGATGATGATGATGATG 1380  
 DB 1321 GGGAGTCTGAATTAATGCAAGAGTGTGATGATGATGATGATGATGATGATGATGATG 1380  
 QY 1381 AATATTCTGTTCTTCTGAGAAATAGACTTCTGAGAGATGATGATGATGATGATGATG 1440  
 DB 1381 AATATTCTGTTCTTCTGAGAAATAGACTTCTGAGAGATGATGATGATGATGATGATG 1440  
 QY 1441 TATGTAAAGTGAAGAGTTCACCTCAATGATGATGATGATGATGATGATGATGATG 1500  
 DB 1441 TATGTAAAGTGAAGAGTTCACCTCAATGATGATGATGATGATGATGATGATGATG 1500  
 QY 1501 TTGGGAAAGCTTATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560  
 DB 1501 TTGGGAAAGCTTATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560  
 QY 1561 TAATTATAGAGAGATTTGTTACTGAGCCACAGATATACAAAGAGCGTCCCTCACAAATA 1620





OY	3781	AAGACCTTCCCGCTCCACACCTGTAATTTGGTAAGTAACAAATACCTCTCAGT	3840
Db	3781	AAGACCTTCCCGCTCCACACCTGTAATTTGGTAAGTAACAAATACCTCTCAGT	3840
OY	3841	CTACTAGGCATAGACACCGTTGCTACCGAGTGTCTGTCTTAAGAACACAGAGGAAATTTAT	3900
Db	3841	CTACTAGGCATAGACACCGTTGCTACCGAGTGTCTGTCTTAAGAACACAGAGGAAATTTAT	3900
OY	3901	TATCATTTGAAGAAATAGCTTAAATGACTGCAGTAACCCAGTAATATTTGGCAAAAGCATCTC	3960
Db	3901	TATCATTTGAAGAAATAGCTTAAATGACTGCAGTAACCCAGTAATATTTGGCAAAAGCATCTC	3960
OY	3961	AGGAACATCACCTTAGTAGAGAAACAAATGTCTGTAGCTGTGTTTCTTACAGTGC	4020
Db	3961	AGGAACATCACCTTAGTAGAGAAACAAATGTCTGTAGCTGTGTTTCTTACAGTGC	4020
OY	4021	GTGAATTTGGAAGACTTGACTGCAATATACAAACCCAGATCTTCTTGTATTTGGTCTT	4080
Db	4021	GTGAATTTGGAAGACTTGACTGCAATATACAAACCCAGATCTTCTTGTATTTGGTCTT	4080
OY	4081	CCAAACAATGAGGCATCAGTCTGTAACCCAGGGAGTTGGTCTGTAGTACAAAGAAATTTGG	4140
Db	4081	CCAAACAATGAGGCATCAGTCTGTAACCCAGGGAGTTGGTCTGTAGTACAAAGAAATTTGG	4140
OY	4141	TTTCAGATGATGAAGAAAGAGAACGGGCTTGGAAAGAAATATTCACAGAGCAACCA	4200
Db	4141	TTTCAGATGATGAAGAAAGAGAACGGGCTTGGAAAGAAATATTCACAGAGCAACCA	4200
OY	4201	TGGATTTCAACTTAGTGAAGAGCATCTGGGTGTGAAGTGAAGAAACAGCTCTCTGAG	4260
Db	4201	TGGATTTCAACTTAGTGAAGAGCATCTGGGTGTGAAGTGAAGAAACAGCTCTCTGAG	4260
OY	4261	ACTGCTCAGGGCTATCTCTCAGAGTGAATTTTAAACATCTCAGCAGAGGGATACCATGC	4320
Db	4261	ACTGCTCAGGGCTATCTCTCAGAGTGAATTTTAAACATCTCAGCAGAGGGATACCATGC	4320
OY	4321	AACATTAACCTGTATAAGCTCCAGCAGGAAATTTGCTGAACCTGTGTTAGAACAC	4380
Db	4321	AACATTAACCTGTATAAGCTCCAGCAGGAAATTTGCTGAACCTGTGTTAGAACAC	4380
OY	4381	ATGGAGAGCAGCCTTCTAACAGATCACTTCATCATATAGAGACCTCTCCCTTGAG	4440
Db	4381	ATGGAGAGCAGCCTTCTAACAGATCACTTCATCATATAGAGACCTCTCCCTTGAG	4440
OY	4441	ACCTGCGAATTCACAGACAAAGCAGATCAGAAAAAGCAGTATTAACTTCACAGAAAA	4500
Db	4441	ACCTGCGAATTCACAGACAAAGCAGATCAGAAAAAGCAGTATTAACTTCACAGAAAA	4500
OY	4501	GTGAATACCCATTAAGCCAGAAATCCAGAAAGGCTTTCTGCTGACAAAGTTTAGTGTCTG	4560
Db	4501	GTGAATACCCATTAAGCCAGAAATCCAGAAAGGCTTTCTGCTGACAAAGTTTAGTGTCTG	4560
OY	4561	CAGATAGTTCTACACAGTAAAAATTAAGAACCCAGAGTAGTGAAGGTCAATCCCTCTTAAT	4620
Db	4561	CAGATAGTTCTACACAGTAAAAATTAAGAACCCAGAGTAGTGAAGGTCAATCCCTCTTAAT	4620
OY	4621	GCCCATCATTTAGATGATAGTGTGTACATGCACAGTTGCTCTGGGAGAGTTCAGAAATGAA	4680
Db	4621	GCCCATCATTTAGATGATAGTGTGTACATGCACAGTTGCTCTGGGAGAGTTCAGAAATGAA	4680
OY	4681	ACTAACCCTTCACAGAGAGGCTCAATTAAGTTGTTGATGTGAGAGCAACAGCTGGAG	4740
Db	4681	ACTAACCCTTCACAGAGAGGCTCAATTAAGTTGTTGATGTGAGAGCAACAGCTGGAG	4740
OY	4741	AGTCTGGGCCACAGATTTGAACGGAACATCTTACTTGGCCAAAGCAGATCTAGAGGAA	4800
Db	4741	AGTCTGGGCCACAGATTTGAACGGAACATCTTACTTGGCCAAAGCAGATCTAGAGGAA	4800
OY	4801	CCCTTACCTTGAATCTGGGATTCAGCCTCTTCTCTGATGACCCCTGAATCTGTATCTCTG	4860
Db	4801	CCCTTACCTTGAATCTGGGATTCAGCCTCTTCTCTGATGACCCCTGAATCTGTATCTCTG	4860

QY	4861	AAGACAGAGCCCCACAGACAGTCTGTTGGCAATATCCATCTTCAACCTCTGATTTGA	4920
Db	4861	AAGACAGAGCCCCACAGAGTCACTGCTGTTGGCAATATCCATCTTCAACCTCTGATTTGA	4920
QY	4921	AAGTTCCCAATTGAAAGTTGCAGATCTGCCAGAGTCCAGCTGCTCATACTCTG	4980
Db	4921	AAGTTCCCAATTGAAAGTTGCGAATCTGCCAGAGTCCAGCTGCTCATACTCTG	4980
QY	4981	ATACTGCGGGGATATATGTCAMTGGAAAGATGAGAGGAGAGAGGACAAATTGACAG	5040
Db	4981	ATACTGCGGGGATATATGTCAMTGGAAAGATGAGAGGAGAGGAGGACAAATTGACAG	5040
QY	5041	CTTCACAGAAAAGGGCTCAACAAAAGAAATGTCATGTTGCTGCGCCTGACCCAGAG	5100
Db	5041	CTTCACAGAAAAGGGCTCAACAAAAGAAATGTCATGTTGCTGCGCCCTGACCCAGAG	5100
QY	5101	AATTTATCTCTGTATGACAAAGTTTGGCAGAAAAACCAATACATTAACTATATTA	5160
Db	5101	AATTTATCTCTGTATGACAAAGTTTGGCAGAAAAACCAATACATTAACTATATTA	5160
QY	5161	CTGAGAGACTACTCATCTTGTATTATGAAGAACAGATGCTGAGTGTGTGTGACAGGACAC	5220
Db	5161	CTGAGAGACTACTCATCTTGTATTATGAAGAACAGATGCTGAGTGTGTGTGACAGGACAC	5220
QY	5221	TGAATATTTTCTAGGAATTGGGGAGAGAAATGGGTATGTTAGCTATTTCTGGGTGACC	5280
Db	5221	TGAATATTTTCTAGGAATTGGGGAGAGAAATGGGTATGTTAGCTATTTCTGGGTGACC	5280
QY	5281	AGTCATTTAAAGAAAGAAATATGCGATGTGAGCATGATTTGAAAGTCAGAGAGATGG	5340
Db	5281	AGTCATTTAAAGAAAGAAATATGCGATGTGAGCATGATTTGAAAGTCAGAGAGATGG	5340
QY	5341	TCAATGAGAAAGAACCCAGAGTCCAAAGCAGAGCAAGAAATCCAGAGACAGAAAGATCT	5400
Db	5341	TCAATGAGAAAGAACCCAGAGTCCAAAGCAGAGCAAGAAATCCAGAGACAGAAAGATCT	5400
QY	5401	TCAGAGGGCTTGAAATCTGTTGCTATGGGCCCCCTTACCACATATGCCACATCAACTGG	5460
Db	5401	TCAGAGGGCTTGAAATCTGTTGCTATGGGCCCCCTTACCACATATGCCACATCAACTGG	5460
QY	5461	AATGATGTGACAGCTGTGTGGTGTCTGTGTGAGTGAAGAGACTTTCAATCACTTACCTTGG	5520
Db	5461	AATGATGTGACAGCTGTGTGGTGTCTGTGTGAGTGAAGAGACTTTCAATCACTTACCTTGG	5520
QY	5521	GCACAGCTGTCCACCAATTGCTGTGTTGCGACGACGATGCCCTGGACACAGCAATGGCT	5580
Db	5521	GCACAGCTGTCCACCAATTGCTGTGTTGCGACGACGATGCCCTGGACACAGCAATGGCT	5580
QY	5581	TCCATATGCAATTTGGGAGATGTGTGAGCACTGTGTGTATCCCGAGAGTGGTGTGGACA	5640
Db	5581	TCCATATGCAATTTGGGAGATGTGTGAGCACTGTGTGTATCCCGAGAGTGGTGTGGACA	5640
QY	5641	GTGTAGCACTTACCAAGTCCAGAGAGCTGACACCTACCTATATACCCAGATCCCCACA	5700
Db	5641	GTGTAGCACTTACCAAGTCCAGAGAGCTGACACCTACCTATATACCCAGATCCCCACA	5700
QY	5701	GCCACTACTGA 5711	
Db	5701	GCCACTACTGA 5711	

US-09-982-828-3  
RESULT 4  
Sequence 3, Application US/09982828  
Publication No. US20030022184A1  
GENERAL INFORMATION:  
APPLICANT: Murphy, Patricia D.  
Allen, Antonette C.  
Critz, Brenda S.  
Olson, Sheri J.  
Thunder, Denise  
Zeng, Bin



```

? TITLE OF INVENTION: Coding Sequences of the Human
? BRCA1 Gene
? NUMBER OF SEQUENCES: 72
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Morgan Lewis & Bockius LLP
? STREET: 1111 Pennsylvania Avenue N. W.
? CITY: Washington
? STATE: DC
? COUNTRY: USA
? ZIP: 20004
?
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent Release #1.0, Version #1.30
?
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/982,828
? FILING DATE: 22-Oct-2001
? CLASSIFICATION: <Unknown>
?
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: US 09/074,453
? FILING DATE: 1998-05-06
? APPLICATION NUMBER: US 08/798,691
? FILING DATE: 1997-02-12
? APPLICATION NUMBER: US 08/598,591
? FILING DATE: 1996-02-12
?
? ATTORNEY/AGENT INFORMATION:
? NAME: Michael S. Tuscan
? REGISTRATION NUMBER: 43,210
? REFERENCE/DOCKET NUMBER: 44921-5053-01-US
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 202-739-3000
? TELEFAX: 202-739-3001
?
? INFORMATION FOR SEQ ID NO: 3:
?
? SEQUENCE CHARACTERISTICS:
? LENGTH: 5711 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: not relevant
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? ORIGINAL SOURCE:
? ORGANISM: Homo sapiens
? STRAIN: BRCA1 (om12)
? POSITION IN GENOME:
? CHROMOSOME/SEGMENT: 17
? MAP POSITION: 17q21
? SEQUENCE DESCRIPTION: SEQ ID NO: 3:
?
? US-09-982-828-3
?
? Query Match 99.9%; Score 5704.6; DB 9; Length 5711;
? Best Local Similarity 99.9%; Pred. No. 0;
? Matches 5707; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 301 GTCCCTTATGTAAATGATATTAACCAAAAGAGCTACAGAAAGTACGAGATTAGTC 360
DB |||||||
QY 301 GTCCCTTATGTAAATGATATTAACCAAAAGAGCTACAGAAAGTACGAGATTAGTC 360
DB |||||||
QY 361 AACTTTGTAAGAGCTATTTGAAATTCATTGTGCTTTACGCTTGACACAGGTTTGAGT 420
DB |||||||
QY 361 AACTTTGTAAGAGCTATTTGAAATTCATTGTGCTTTACGCTTGACACAGGTTTGAGT 420
DB |||||||
QY 421 ATGCAAACAGCTATTAATTTTGCAGAAAAGAAATTAATCTCTCGAATCATCTAAAGATG 480
DB |||||||
QY 421 ATGCAAACAGCTATTAATTTTGCAGAAAAGAAATTAATCTCTCGAATCATCTAAAGATG 480
DB |||||||
QY 481 AAGTTTCTATCATCCAAAGATATGGGCTACAGAAACCGTCCAAAGACTTTACAGAGTG 540
DB |||||||
QY 481 AAGTTTCTATCATCCAAAGATATGGGCTACAGAAACCGTCCAAAGACTTTACAGAGTG 540
DB |||||||
QY 541 AACCCGAAATCCCTTCCCTGAGGAAACCACTCTCAGTCCAACTCTAACCTTGGA 600
DB |||||||
QY 541 CTTGAGAGACCTGAGGACCAAGGAGGATACCAACCTCAAGACGCTGCTACATTTG 660
DB |||||||
QY 601 CTTGAGAGACCTGAGGACCAAGGAGGATACCAACCTCAAGACGCTGCTACATTTG 660
DB |||||||
QY 601 CTTGAGAGACCTGAGGACCAAGGAGGATACCAACCTCAAGACGCTGCTACATTTG 660
DB |||||||
QY 661 AATTGGGATCGATTCTTCTGAAAGATACCGTTAATAAGCACTTATTCAGTGGGAG 720
DB |||||||
QY 661 AATTGGGATCGATTCTTCTGAAAGATACCGTTAATAAGCACTTATTCAGTGGGAG 720
DB |||||||
QY 721 ATCAAGAAATGTTTACAAATCACCCCTCAAGSACAGGATGAATACGTTTGGATTCG 780
DB |||||||
QY 721 ATCAAGAAATGTTTACAAATCACCCCTCAAGSACAGGATGAATACGTTTGGATTCG 780
DB |||||||
QY 781 CAAAAAGGCTGCTTGAATTTTCTGAGACGATGTAACAAATCTGAATCATATCAAC 840
DB |||||||
QY 781 CAAAAAGGCTGCTTGAATTTTCTGAGACGATGTAACAAATCTGAATCATATCAAC 840
DB |||||||
QY 841 CCAATAATTAATGATTTGACACCACTGAGAGGCTGACGAGAGGCTACAGAAAGT 900
DB |||||||
QY 841 CCAATAATTAATGATTTGACACCACTGAGAGGCTGACGAGAGGCTACAGAAAGT 900
DB |||||||
QY 901 ATCAAGGATGTTCTGTTTCAAACTTGCATGTGAGCCATGTGGCAAAATCTCATGCA 960
DB |||||||
QY 901 ATCAAGGATGTTCTGTTTCAAACTTGCATGTGAGCCATGTGGCAAAATCTCATGCA 960
DB |||||||
QY 961 GCTATTACAGCAGAGACGACGAGTTTATTAATCACTTAAGACAGATGAATGTAGAA 1020
DB |||||||
QY 961 GCTATTACAGCAGAGACGACGAGTTTATTAATCACTTAAGACAGATGAATGTAGAA 1020
DB |||||||
QY 1021 AGCCTCAATTTCTGTAATAAAGCAACAGCCTGCTTAGCAAGGAGCCACATTAACAGAT 1080
DB |||||||
QY 1021 AGCCTCAATTTCTGTAATAAAGCAACAGCCTGCTTAGCAAGGAGCCACATTAACAGAT 1080
DB |||||||
QY 1081 GGGCTGCAAGTAAAGCAATGTATGATAGCGGACTCCGACAGCAAGAAAAAGGTAG 1140
DB |||||||
QY 1081 GGGCTGCAAGTAAAGCAATGTATGATAGCGGACTCCGACAGCAAGAAAAAGGTAG 1140
DB |||||||
QY 1141 ATTCGAATGCTGATCCCTGCTGTGAGAGAAAGATGAATGAACAGAACTGCATGCT 1200
DB |||||||
QY 1141 ATTCGAATGCTGATCCCTGCTGTGAGAGAAAGATGAATGAACAGAACTGCATGCT 1200
DB |||||||
QY 1201 CAGAGAAATCTAGAGATACGTAAGATGTTCTCTGATTAACCTAATAGAGCAATTCAGA 1260
DB |||||||
QY 1201 CAGAGAAATCTAGAGATACGTAAGATGTTCTCTGATTAACCTAATAGAGCAATTCAGA 1260
DB |||||||
QY 1261 AAGTTAATGAGTGTTCAGAGAGTGAATGATGAGTTAGGTTCTGATGACTACATGATG 1320
DB |||||||
QY 1261 AAGTTAATGAGTGTTCAGAGAGTGAATGATGAGTTAGGTTCTGATGACTACATGATG 1320
DB |||||||
QY 1321 GGGAGCTGGAATCAAAAGTCAAAAGTACGTAATGATGAGAGTCTAAATAGGTAGATG 1380
DB |||||||
QY 1321 GGGAGCTGGAATCAAAAGTCAAAAGTACGTAATGATGAGAGTCTAAATAGGTAGATG 1380
DB |||||||
QY 1381 AATATCTGTTCTTCAGAGAAATATAGACTTAGGCGCAGTATCTCATGAGGCTTTAA 1440
DB |||||||

```

Db	1381	AAATATTCCTGTTCTTCAGAGAAATATGACTTACGCGCAGTGAATCCATATGAGGCTTTAA	1440
Qy	1441	TATGTAAAAGTGAAGAGTTCACCTCAATCAGTAGAGAGTAATATTGAACAAATATAT	1500
Db	1441	TATGTAAAAGTGAAGAGTTCACCTCAATCAGTAGAGAGTAATATTGAACAAATATAT	1500
Qy	1501	TTGGGAAAAACCTATCGGAAGAAGCAGCCTCCCACTTAAGCCATGTAACTGAAAAATC	1560
Db	1501	TTGGGAAAAACCTATCGGAAGAAGCAGCCTCCCACTTAAGCCATGTAACTGAAAAATC	1560
Qy	1561	TAAATTATAGAGACATTTGTACTGAGCCACAGATATATCAAGAGCGTCCCTCCACAATA	1620
Db	1561	TAAATTATAGAGACATTTGTACTGAGCCACAGATATATCAAGAGCGTCCCTCCACAATA	1620
Qy	1621	AATTTAAACGCTTAAAGAGACCTACATAGGCTTCATCCTGAGAGATTTTTCAGAAG	1680
Db	1621	AATTTAAACGCTTAAAGAGACCTACATAGGCTTCATCCTGAGAGATTTTTCAGAAG	1680
Qy	1681	CAGATTTGGCAGCTTCAAAAAGACTCCTGAAATGATTAATCAGGAACTTAACCAAGCGAGC	1740
Db	1681	CAGATTTGGCAGCTTCAAAAAGACTCCTGAAATGATTAATCAGGAACTTAACCAAGCGAGC	1740
Qy	1741	AGAATGTCAAGTGTATGTAATATTACTAATAGTGTGATGAGAAATTAACCAAGAGTAT	1800
Db	1741	AGAATGTCAAGTGTATGTAATATTACTAATAGTGTGATGAGAAATTAACCAAGAGTAT	1800
Qy	1801	CTATTTCAGAAATGAGAAAAATCCTTAACCAATAGAAATCTGAAAAAGAAATCTGCTTCA	1860
Db	1801	CTATTTCAGAAATGAGAAAAATCCTTAACCAATAGAAATCTGAAAAAGAAATCTGCTTCA	1860
Qy	1861	AAACGAAAGCTGAACCTATTAAGCAGCAGATTAAGCAATATGGAATCGAATTTAAATATCC	1920
Db	1861	AAACGAAAGCTGAACCTATTAAGCAGCAGATTAAGCAATATGGAATCGAATTTAAATATCC	1920
Qy	1921	ACAATTCAAAAGCACTTAAGAAATAGGCTGAGAGAGAACTTTCTTACAGGCAATTTTC	1980
Db	1921	ACAATTCAAAAGCACTTAAGAAATAGGCTGAGAGAGAACTTTCTTACAGGCAATTTTC	1980
Qy	1981	ATGCGCTTGAAGTACTAGTACAGTAAATCTTAAGCCCACTAATTTGTAAGTAAATGGA	2040
Db	1981	ATGCGCTTGAAGTACTAGTACAGTAAATCTTAAGCCCACTAATTTGTAAGTAAATGGA	2040
Qy	2041	TTGATAGTGTCTTACAGCAGTAAGAGATTAAGAAAAAAAGTACACCAATATGCGACATCA	2100
Db	2041	TTGATAGTGTCTTACAGCAGTAAGAGATTAAGAAAAAAAGTACACCAATATGCGACATCA	2100
Qy	2101	GGCAGAGCAAACTTAACACTCATGTGAAAGTAAAGAACTGCAACTGAGCCCAAGAGA	2160
Db	2101	GGCAGAGCAAACTTAACACTCATGTGAAAGTAAAGAACTGCAACTGAGCCCAAGAGA	2160
Qy	2161	GTAACAAAGCAAAATGACAGACAGTAAGAAACATGACAGTACTTTCCACAGAGCTGA	2220
Db	2161	GTAACAAAGCAAAATGACAGACAGTAAGAAACATGACAGTACTTTCCACAGAGCTGA	2220
Qy	2221	AGTTAAACAATATGACACCTGTGTTCTTTTACTAAGTGTCCAAATATCCAGTGAATTAAGAAAT	2280
Db	2221	AGTTAAACAATATGACACCTGTGTTCTTTTACTAAGTGTCCAAATATCCAGTGAATTAAGAAAT	2280
Qy	2281	TTTGTCAATCTTAGCCTTCCAAAGAGAAAGAAAGAACTTGAACACGTTTAAAGTGT	2340
Db	2281	TTTGTCAATCTTAGCCTTCCAAAGAGAAAGAAAGAACTTGAACACGTTTAAAGTGT	2340
Qy	2341	CTAATATATGCTGAAGAACCCCAAAAGATGATCAATGTAAGTGAAGAGGTTTTCGAACCTG	2400
Db	2341	CTAATATATGCTGAAGAACCCCAAAAGATGATCAATGTAAGTGAAGAGGTTTTCGAACCTG	2400
Qy	2401	AAAGATCTGTAGAGAGTATTTTCAATGTGTATCTGTAGCTGATTTATGGAACCTAGG	2460
Db	2401	AAAGATCTGTAGAGAGTATTTTCAATGTGTATCTGTAGCTGATTTATGGAACCTAGG	2460
Qy	2461	AAAGTATCTGTTACTGGAAGTATGACCTTAGGGAAGGCAAAAAACGAACCAATTAAT	2520
Db	2461	AAAGTATCTGTTACTGGAAGTATGACCTTAGGGAAGGCAAAAAACGAACCAATTAAT	2520
Qy	2521	GTGTAGAGTCAGTGTGACAGCATTTTGAAGAACCCCAAGGGCAATTAATCAATGTGTTCCAAAG	2580
Db	2521	GTGTAGAGTCAGTGTGACAGCATTTTGAAGAACCCCAAGGGCAATTAATCAATGTGTTCCAAAG	2580
Qy	2581	ATAATAGAAATGACACAGAGGCTTTTAAGTATCCATTGGACATGAGATTAAACACAGCTC	2640
Db	2581	ATAATAGAAATGACACAGAGGCTTTTAAGTATCCATTGGACATGAGATTAAACACAGCTC	2640
Qy	2641	GGGAAACAAGCATTAATGGAAGAAAGTGAACCTGTATGTCGTATTTTGCAGATATCAT	2700
Db	2641	GGGAAACAAGCATTAATGGAAGAAAGTGAACCTGTATGTCGTATTTTGCAGATATCAT	2700
Qy	2701	TCAGAGTTTCAAGAGCCAGTCAATTTCTCTGTGTTTCAATTCAGAGAAATTCAGAGAGAG	2760
Db	2701	TCAGAGTTTCAAGAGCCAGTCAATTTCTCTGTGTTTCAATTCAGAGAAATTCAGAGAGAG	2760
Qy	2761	AATGTGCAACATTTCTGCCACATCTGGGTCTTTAAGAAACAAGTCCAAAGTCACTT	2820
Db	2761	AATGTGCAACATTTCTGCCACATCTGGGTCTTTAAGAAACAAGTCCAAAGTCACTT	2820
Qy	2821	TTGAATGTGAACAAAGAGAAATATCAAGAAAGAAATGAGTGAATTAATCAAGCCCTGAC	2880
Db	2821	TTGAATGTGAACAAAGAGAAATATCAAGAAAGAAATGAGTGAATTAATCAAGCCCTGAC	2880
Qy	2881	AGACAGTTAATATCAGTCAAGGCTTCTGCTGTGGTGGTGAGAAAGATTAAGCAGTTGATA	2940
Db	2881	AGACAGTTAATATCAGTCAAGGCTTCTGCTGTGGTGGTGAGAAAGATTAAGCAGTTGATA	2940
Qy	2941	ATGCCAAATGTATGATTCAGAGAGGCTGTAGGTTTGTCTATCATCTCAGTTTCAGAGCA	3000
Db	2941	ATGCCAAATGTATGATTCAGAGAGGCTGTAGGTTTGTCTATCATCTCAGTTTCAGAGCA	3000
Qy	3001	ACGAACATGAGCATCTATTACTCCAAATTAACATGAGCTTTTACAAACCCATATGCTATAC	3060
Db	3001	ACGAACATGAGCATCTATTACTCCAAATTAACATGAGCTTTTACAAACCCATATGCTATAC	3060
Qy	3061	CACCACTTTTCCCATCAAGTCAATTTGTTTAAACTTAAGTAAAGAAATTCGCTAGAGG	3120
Db	3061	CACCACTTTTCCCATCAAGTCAATTTGTTTAAACTTAAGTAAAGAAATTCGCTAGAGG	3120
Qy	3121	AAAACCTTTGAGGAACATTCATATGTCACCTGAAGAGAAATGGGAAATGGAACATTTCCA	3180
Db	3121	AAAACCTTTGAGGAACATTCATATGTCACCTGAAGAGAAATGGGAAATGGAACATTTCCA	3180
Qy	3181	GTCAGTGAAGCAAAATTAAGCCGTATTAACATTTGAGAAAAATGTTTTTAAAGAGCCACT	3240
Db	3181	GTCAGTGAAGCAAAATTAAGCCGTATTAACATTTGAGAAAAATGTTTTTAAAGAGCCACT	3240
Qy	3241	CAAGCAATATTAATGAAGTAGGTTCCAGTACTAATGAAGTGGGCTCCAGTATATTAATGA	3300
Db	3241	CAAGCAATATTAATGAAGTAGGTTCCAGTACTAATGAAGTGGGCTCCAGTATATTAATGA	3300
Qy	3301	TAGGTTCCAGTGATGAAGAACATTCACAGACAGTAGGTAGAAACAGAGGGCCAAATTTGA	3360
Db	3301	TAGGTTCCAGTGATGAAGAACATTCACAGACAGTAGGTAGAAACAGAGGGCCAAATTTGA	3360
Qy	3361	ATGCTATCTTATAGATTAGGGGTTTTCGACACTAGTCTATTAACAAAGTCTTCTCGAA	3420
Db	3361	ATGCTATCTTATAGATTAGGGGTTTTCGACACTAGTCTATTAACAAAGTCTTCTCGAA	3420
Qy	3421	GTAATGTATAGCATCTGGAATTAAGAAAGCAAGATATGAAGAGTACGCTGACTGTA	3480
Db	3421	GTAATGTATAGCATCTGGAATTAAGAAAGCAAGATATGAAGAGTACGCTGACTGTA	3480
Qy	3481	ATACAGATTTCTCCATATCTGATTTTCAATTAACATTAAGTGAAGAGTATGGAAGTATC	3540
Db	3481	ATACAGATTTCTCCATATCTGATTTTCAATTAACATTAAGTGAAGAGTATGGAAGTATC	3540
Qy	3541	ATGCATCTCAGGTTTGTCTTGAGACACCTGATGACCTGTTAGATGATGGTGAATTAAGG	3600
Db	3541	ATGCATCTCAGGTTTGTCTTGAGACACCTGATGACCTGTTAGATGATGGTGAATTAAGG	3600

OY 3601 AAGATCTAGTTTGGTGAATAATGACATTAAAGGAAGTTCTGCTGTTTTTAGCAAAACGC 3660  
|||||  
Db 3601 AAGATCTAGTTTGGTGAATAATGACATTAAAGGAAGTTCTGCTGTTTTTAGCAAAACGC 3660  
OY 3661 TCCGAAAGAGAGCTTAGAGAGAGTCTAGCCCTTCCACCATPACATTTGGCTCAGG 3720  
|||||  
Db 3661 TCCGAGAGAGAGCTTAGAGAGAGTCTAGCCCTTCCACCATPACATTTGGCTCAGG 3720  
OY 3721 GTTCCGAGAGAGGGGCCAAGAAATTAGAGTCCAGAGAAAGACTTACTAGTGAAGTGG 3780  
|||||  
Db 3721 GTTCCGAGAGAGGGGCCAAGAAATTAGAGTCCAGAGAAAGACTTACTAGTGAAGTGG 3780  
OY 3781 AAGAGCTTCCCTGCTCCAACTTGTATTGGTAAAGTAACATATACCTTCTCAGT 3840  
|||||  
Db 3781 AAGAGCTTCCCTGCTCCAACTTGTATTGGTAAAGTAACATATATACCTTCTCAGT 3840  
OY 3841 CTACTAGGATAGACACCGTCTGCTACCGAGTGTCTGTAGAGACACAGAGAGAAATTTAT 3900  
|||||  
Db 3841 CTACTAGGATAGACACCGTCTGCTACCGAGTGTCTGTAGAGACACAGAGAGAAATTTAT 3900  
OY 3901 TATCATTTGAAGAATAGCTTAATGACTGCTAATACAGAGTAATTTGGCAAGGATCTC 3960  
|||||  
Db 3901 TATCATTTGAAGAATAGCTTAATGACTGCTAATACAGAGTAATTTGGCAAGGATCTC 3960  
OY 3961 AGGAACATCACCTTAGTGAAGAAACAAATGTCTGCTAGCTGTTTCTTCCACAGTSCA 4020  
|||||  
Db 3961 AGGAACATCACCTTAGTGAAGAAACAAATGTCTGCTAGCTGTTTCTTCCACAGTSCA 4020  
OY 4021 GTGATTTGGAGACTTCTGCTCAATPACAAACCCAGGATCTTCTGTAATGGTCTT 4080  
|||||  
Db 4021 GTGATTTGGAGACTTCTGCTCAATPACAAACCCAGGATCTTCTGTAATGGTCTT 4080  
OY 4081 CCAACCAATAGGATAGTCTGAAGCCAGGGAGTGTCTGAGTACAGAGAAATGG 4140  
|||||  
Db 4081 CCAACCAATAGGATAGTCTGAAGCCAGGGAGTGTCTGAGTACAGAGAAATGG 4140  
OY 4141 TTTTCAGATGATGAAGAAAGAGAAACGGCTTGAAGAAATATCAAGAAAGCA 4200  
|||||  
Db 4141 TTTTCAGATGATGAAGAAAGAGAAACGGCTTGAAGAAATATCAAGAAAGCA 4200  
OY 4201 TGGATTTCAAACCTTGGTGAAGCAGCATCTGGGTGAGAGTGAACAAGCCCTCTGAG 4260  
|||||  
Db 4201 TGGATTTCAAACCTTGGTGAAGCAGCATCTGGGTGAGAGTGAACAAGCCCTCTGAG 4260  
OY 4261 ACTGCTAGGGCTATCTCTCAGAGTGAATTTAACCATCCAGAGAGGATACCATGC 4320  
|||||  
Db 4261 ACTGCTAGGGCTATCTCTCAGAGTGAATTTAACCATCCAGAGAGGATACCATGC 4320  
OY 4321 AACATTAACCTGATAAGCTCCAGCAGAAATGCTGTAAGTGAAGCTGTGTAGAACAGC 4380  
|||||  
Db 4321 AACATTAACCTGATAAGCTCCAGCAGAAATGCTGTAAGTGAAGCTGTGTAGAACAGC 4380  
OY 4381 ATGGAGAGCCAGCTTCTTAACAGCTACCTTCCATATATAGTACTTTCGCCCTTGAGG 4440  
|||||  
Db 4381 ATGGAGAGCCAGCTTCTTAACAGCTACCTTCCATATATAGTACTTTCGCCCTTGAGG 4440  
OY 4441 ACCTGCGAAATCCAGAACCAAGCAATCAGAAAAAGCAAGTATTAACCTTCAAGAAAAATA 4500  
|||||  
Db 4441 ACCTGCGAAATCCAGAACCAAGCAATCAGAAAAAGCAAGTATTAACCTTCAAGAAAAATA 4500  
OY 4501 GTGAATACCTTAAGCCAGAAATCCAGAAAGCCCTTTCGTGACAAAGTTTGAAGTGTCTG 4560  
|||||  
Db 4501 GTGAATACCTTAAGCCAGAAATCCAGAAAGCCCTTTCGTGACAAAGTTTGAAGTGTCTG 4560  
OY 4561 CAGTAGTCTTACCAAGTAAATTAAGAACCCAGAGAGTGAAGAGTCAATCCCTTCTAAT 4620  
|||||  
Db 4561 CAGTAGTCTTACCAAGTAAATTAAGAACCCAGAGAGTGAAGAGTCAATCCCTTCTAAT 4620  
OY 4621 GCCCATCATTTAGATGATAGGTGTACATGACAGAGTGTCTGGAGTCTTCAAGATAGAA 4680  
|||||  
Db 4621 GCCCATCATTTAGATGATAGGTGTACATGACAGAGTGTCTGGAGTCTTCAAGATAGAA 4680

OY 4681 ACTACCATCTCAAGAGAGCTCATTAAGGTGTTGATGTGGAGGACCAACGCTGAG 4740  
|||||  
Db 4681 ACTACCATCTCAAGAGAGCTCATTAAGGTGTTGATGTGGAGGACCAACGCTGAG 4740  
OY 4741 AGTGGGCCACAGATTTGACGGAACAATCTTACTTCCCAAGGCAAGATCTAGAGGAA 4800  
|||||  
Db 4741 AGTGGGCCACAGATTTGACGGAACAATCTTACTTCCCAAGGCAAGATCTAGAGGAA 4800  
OY 4801 CCCCTTACTGGAATCTGGAATCAGGCTCTCTGATGAGACCCCTGAATCTGATCTCTG 4860  
|||||  
Db 4801 CCCCTTACTGGAATCTGGAATCAGGCTCTCTGATGAGACCCCTGAATCTGATCTCTG 4860  
OY 4861 AAGACAGAGCCCAAGAGTACGCTGTGTGGCAATACATCTTCAACCTCTGATTTGA 4920  
|||||  
Db 4861 AAGACAGAGCCCAAGAGTACGCTGTGTGGCAATACATCTTCAACCTCTGATTTGA 4920  
OY 4921 AAGTTCCCAATTTGAAGTTGCAGAAATCTGCCAGATCCAGCTGCTCATACTAGT 4980  
|||||  
Db 4921 AAGTTCCCAATTTGAAGTTGCAGAAATCTGCCAGATCCAGCTGCTCATACTAGT 4980  
OY 4981 ATACTGCTGGGTATATATGCAATGGAAGAAAGTGTGAGCAGGAGAGACCAATTTGACAG 5040  
|||||  
Db 4981 ATACTGCTGGGTATATATGCAATGGAAGAAAGTGTGAGCAGGAGAGAAAGCAATTTGACAG 5040  
OY 5041 CTTCACAGAAAGGCTCAACAAAAGATGTCATGCTGTCTGCTGAGCCCTGAG 5100  
|||||  
Db 5041 CTTCACAGAAAGGCTCAACAAAAGATGTCATGCTGTCTGCTGAGCCCTGAGCCCTGAG 5100  
OY 5101 AATTTATGCTGCTGATACAGTTTGGCAGAAAACACCATCATCTTAACTAATTA 5160  
|||||  
Db 5101 AATTTATGCTGCTGATACAGTTTGGCAGAAAACACCATCATCTTAACTAATTA 5160  
OY 5161 CTGAAGAGACTACCTATGTTGTTATGAAGACAGATGCTGATGTTGTGTGAGAGGAC 5220  
|||||  
Db 5161 CTGAAGAGACTACCTATGTTGTTATGAAGACAGATGCTGATGTTGTGTGAGAGGAC 5220  
OY 5221 TGAATATTTTCTAGGAATTCGGGAGGAAATGGGTAGTATGCTATTTCTGGGTGACCC 5280  
|||||  
Db 5221 TGAATATTTTCTAGGAATTCGGGAGGAAATGGGTAGTATGCTATTTCTGGGTGACCC 5280  
OY 5281 AGTCTATTTAAAGAAAGAAATGCTGAATGACATGATTTGAAGTCAAGAGAGATG 5340  
|||||  
Db 5281 AGTCTATTTAAAGAAAGAAATGCTGAATGACATGATTTGAAGTCAAGAGAGATG 5340  
OY 5341 TCAATGGAAGAAACCAAGGCTCCAAAGCAGCAAGAGATCCAGAGACGAAGATCT 5400  
|||||  
Db 5341 TCAATGGAAGAAACCAAGGCTCCAAAGCAGCAAGAGATCCAGAGACGAAGATCT 5400  
OY 5401 TCAGGGGCTAGAAATCTGTTGCTATGAGGCCCTTACCAACATGCTCCACAGATCACTGG 5460  
|||||  
Db 5401 TCAGGGGCTAGAAATCTGTTGCTATGAGGCCCTTACCAACATGCTCCACAGATCACTGG 5460  
OY 5461 AATGATGATGACAGTGTGTGCTTCTGTGTGAGAGAGCTTTCATCATTTACCCCTTG 5520  
|||||  
Db 5461 AATGATGATGACAGTGTGTGCTTCTGTGTGAGAGAGCTTTCATCATTTACCCCTTG 5520  
OY 5521 AATGATGATGACAGTGTGTGCTTCTGTGTGAGAGAGCTTTCATCATTTACCCCTTG 5580  
|||||  
Db 5521 AATGATGATGACAGTGTGTGCTTCTGTGTGAGAGAGCTTTCATCATTTACCCCTTG 5580  
OY 5581 GCACAGGTGTCACCCCAATTTGTGTGTGAGCCAGAGATGCTTGAACAAGCAATGCT 5640  
|||||  
Db 5581 GCACAGGTGTCACCCCAATTTGTGTGTGAGCCAGAGATGCTTGAACAAGCAATGCT 5640  
OY 5641 TCCATGCAATTTGGGAGAGATGTTGAGAGCAGCTGTGAGTACCCAGAGAGTGGTGTGGACA 5700  
|||||  
Db 5641 TCCATGCAATTTGGGAGAGATGTTGAGAGCAGCTGTGAGTACCCAGAGAGTGGTGTGGACA 5700  
OY 5701 GTTAGAGACTTACAGAGTGCAGAGAGTGCACACTACTATACCCAGATGCCCAACA 5760  
|||||  
Db 5701 GTTAGAGACTTACAGAGTGCAGAGAGTGCACACTACTATACCCAGATGCCCAACA 5760  
OY 5761 GCCACTACTGA 5711  
|||||  
Db 5761 GCCACTACTGA 5711

RESULT 5  
US-10-022-819-1  
Sequence 1, Application US/10022819  
Publication No. US20030027166A1  
GENERAL INFORMATION:  
APPLICANT: ALLEN, Antonette C. P.  
OLSEN, Sheri J.  
LAWRENCE, Tammy  
ANGELLY, Tracy S.  
RABIN, Mark B.  
TITLE OF INVENTION: CODING SEQUENCE HAPLOTYPE OF THE HUMAN  
BRCA1 GENE  
NUMBER OF SEQUENCES: 67  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morgan Lewis & Bockius LLP  
STREET: 1111 Pennsylvania Avenue  
City: Washington DC  
STATE: District of Columbia  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/022,819  
FILING DATE: 22-Apr-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/074,452  
FILING DATE: 1998-05-06  
ATTORNEY/AGENT INFORMATION:  
NAME: <Unknown>  
REGISTRATION NUMBER: <Unknown>  
REFERENCE/DOCKET NUMBER: 044921-5049-01-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-739-3000  
TELEFAX: 202-739-3001  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5711 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: HOMO SAPIENS  
STRAIN: BRCA1  
HAPLOTYPE: OMI4  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: 17  
MAP POSITION: 17q21  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-10-022-819-1  
Query Match 99.98; Score 5704.6; DB 9; Length 5711;  
Best Local Similarity 99.98; Pred. No. 0;  
Matches 5707; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AGCTGCTGAGACTTCTGACCCCGACACGAGCTGTGGGTTTCTCAGATACTGGCC 60  
DB 1 AGCTGCTGAGACTTCTGACCCCGACACGAGCTGTGGGTTTCTCAGATACTGGCC 60  
QY 61 CCTGGCTCAGAGAGGCTTACCCCTGCTCTGCTGGGTAAGTTCATTGGAAACAGAA 120  
DB 61 CCTGGCTCAGAGAGGCTTACCCCTGCTCTGCTGGGTAAGTTCATTGGAAACAGAA 120  
QY 121 TGGATTATCTGCTCTTCCGCTTGAAGAAGTACAAAATGTCATTATGCTATGCAGAA 180  
DB 121 TGGATTATCTGCTCTTCCGCTTGAAGAAGTACAAAATGTCATTATGCTATGCAGAA 180

DB 121 TGGATTATCTGCTCTTCCGCTTGAAGAAGTACAAAATGTCATTATGCTATGCAGAA 180  
QY 181 TCTTAGAGTGCCATGCTGTGGAGTTGATCAAGAACTGTCTCCAAAGTGGACC 240  
DB 181 TCTTAGAGTGCCATGCTGTGGAGTTGATCAAGAACTGTCTCCAAAGTGGACC 240  
QY 241 ACATATTTTGGCAATTTTGCATGCTGAACCTTCCACCAAGAAAGGGCTTCCAGCT 300  
DB 241 ACATATTTTGGCAATTTTGCATGCTGAACCTTCCACCAAGAAAGGGCTTCCAGCT 300  
QY 301 GTCTTTATGTAGAAATGATATTAACCAAAAGAGCCCTACAGAAATAGAGATTAGTC 360  
DB 301 GTCTTTATGTAGAAATGATATTAACCAAAAGAGCCCTACAGAAATAGAGATTAGTC 360  
QY 361 AACTTGTGAAGACCTATTAATAATCAATTTTGTCTTTCAGCTTGACACAGTTGGAGT 420  
DB 361 AACTTGTGAAGACCTATTAATAATCAATTTTGTCTTTCAGCTTGACACAGTTGGAGT 420  
QY 421 ATGCAAAACAGCTATATTTTGCAGAAAAGGAAATACCTCCCTGAACATCTAAAGATG 480  
DB 421 ATGCAAAACAGCTATATTTTGCAGAAAAGGAAATACCTCCCTGAACATCTAAAGATG 480  
QY 481 AAGTTTCTATCATCCAAAGTATGGGCTACAGAAACCGTGCACAAACCTTACAGAGTG 540  
DB 481 AAGTTTCTATCATCCAAAGTATGGGCTACAGAAACCGTGCACAAACCTTACAGAGTG 540  
QY 541 AACCAGAAATCCCTCTCAGAGAAACCGTGCAGAGTCCAGCTCTAACCTTGGAA 600  
DB 541 AACCAGAAATCCCTCTCAGAGAAACCGTGCAGAGTCCAGCTCTAACCTTGGAA 600  
QY 601 CTGTGAGAACTGTGAGAGCAAAACAGGAGATACAACTCAAAAGAGCTGTCTACATTG 660  
DB 601 CTGTGAGAACTGTGAGAGCAAAACAGGAGATACAACTCAAAAGAGCTGTCTACATTG 660  
QY 661 AATTGGATCTGATCTTCTGGAAGATACCGTTAATAGCACTTATTGCACTGTGGAG 720  
DB 661 AATTGGATCTGATCTTCTGGAAGATACCGTTAATAGCACTTATTGCACTGTGGAG 720  
QY 721 ATCAAGATTTGTTACAAATACCCCTCAAGAACCGAGTGAATACATGTTGGATTG 780  
DB 721 ATCAAGATTTGTTACAAATACCCCTCAAGAACCGAGTGAATACATGTTGGATTG 780  
QY 781 CAAAAAGGCTGCTTGTGAATTTTGTGACAGGATGTAAACAAATCTGAATCATCAAC 840  
DB 781 CAAAAAGGCTGCTTGTGAATTTTGTGACAGGATGTAAACAAATCTGAATCATCAAC 840  
QY 841 CCACTATATATGATTTGAACACCACTGAGAGGCTGACCTAGAGAGCTCCAGAAAGT 900  
DB 841 CCACTATATATGATTTGAACACCACTGAGAGGCTGACCTAGAGAGCTCCAGAAAGT 900  
QY 901 ATCAGGTAAGTCTGTTCAAACTGATGTGAGGCACTGGCAACAAATACATGCA 960  
DB 901 ATCAGGTAAGTCTGTTCAAACTGATGTGAGGCACTGGCAACAAATACATGCA 960  
QY 961 GCTCATTTACAGATGAGAACAGAGCTTTATTACTCACTAAAGACGAATGATGAGAA 1020  
DB 961 GCTCATTTACAGATGAGAACAGAGCTTTATTACTCACTAAAGACGAATGATGAGAA 1020  
QY 1021 AGGCTGAATCTGATTAATAAGCAACAGGCTGGCTTACAGAGAGCCCAACATCAAT 1080  
DB 1021 AGGCTGAATCTGATTAATAAGCAACAGGCTGGCTTACAGAGAGCCCAACATCAAT 1080  
QY 1081 GGGCTGGAAGTGAAGAAACATGTAATGATAGGAGGATCCCGACAGAAAAAGGTAG 1140  
DB 1081 GGGCTGGAAGTGAAGAAACATGTAATGATAGGAGGATCCCGACAGAAAAAGGTAG 1140  
QY 1141 ATCTGAATGCTGATCCCTGTGTGAGAGAAAGAAATGAAGCAAGAAATGCCATGCT 1200  
DB 1141 ATCTGAATGCTGATCCCTGTGTGAGAGAAAGAAATGAAGCAAGAAATGCCATGCT 1200  
QY 1201 CAGAGAAATCTAGAGATAGTGAAGATGTTCTTGGATTAACACTAATATACCATTTGCA 1260  
DB 1201 CAGAGAAATCTAGAGATAGTGAAGATGTTCTTGGATTAACACTAATATACCATTTGCA 1260

OY	1261	AAGTTAAATGAGTGGTTTCCAGAGATGATGAACCTGTTAGTTCGTATGACTCACATGATG	1320
Db	1261	AAGTTAAATGAGTGGTTTCCAGAGATGATGAACCTGTTAGTTCGTATGACTCACATGATG	1320
OY	1321	GGGAGTCGAAATCAAAATGCCAAACTAGCTATGATGAGCGTCTTAATGAGGTAAATG	1380
Db	1321	GGGAGTCGAAATCAAAATGCCAAACTAGCTATGATGAGCGTCTTAATGAGGTAAATG	1380
OY	1381	AATATTCGTGTTCTTCAGAGAAAATPACACTTACTGAGCAGTGAATCCATGAGGCTTTAA	1440
Db	1381	AATATTCGTGTTCTTCAGAGAAAATPACACTTACTGAGCAGTGAATCCATGAGGCTTTAA	1440
OY	1441	TATGTAAAAGTGAAGAAGCTTCATCCCAATCAGTAGACAGTAAATATTGGAAGACAAATAT	1500
Db	1441	TATGTAAAAGTGAAGAAGCTTCATCCCAATCAGTAGAGAGTAATATTGGAAGACAAATAT	1500
OY	1501	TTGGGAAAACCTATCGGAGAGAAGGCAAGCCCTCCCACTTAAGGCATGTAACCTGAATC	1560
Db	1501	TTGGGAAAACCTATCGGAGAGAAGGCAAGCCCTCCCACTTAAGGCATGTAACCTGAATC	1560
OY	1561	TAAATTATAGAGCATTTTGTTACTGAGCCACAGATAATACAAAGCGTCCCTCCACAATA	1620
Db	1561	TAAATTATAGAGCATTTTGTTACTGAGCCACAGATAATACAAAGCGTCCCTCCACAATA	1620
OY	1621	AATTAAGGCGTAAAGGAGACCTACATCAGGCGCTTCACTCCGAGAGATTTTATCAAGAAAG	1680
Db	1621	AATTAAGGCGTAAAGGAGACCTACATCAGGCGCTTCACTCCGAGAGATTTTATCAAGAAAG	1680
OY	1681	CAGATTTGGCAGTTCAAAAAGCATCCTGAAATGATAAATCAGGGAACCTAACCAACGGAGC	1740
Db	1681	CAGATTTGGCAGTTCAAAAAGCATCCTGAAATGATAAATCAGGGAACCTAACCAACGGAGC	1740
OY	1741	AGATGGCTCAAGTATGATATTTCTAATATAGTGTGATGAGATTAATCAAAAGGTATTT	1800
Db	1741	AGATGGCTCAAGTATGATATTTCTAATATAGTGTGATGAGATTAATCAAAAGGTATTT	1800
OY	1801	CTATTTCAGAAATGAGAAAATCTTAACCCAAATAGAAATACTCGAAAAGAAATCTGCTTTCA	1860
Db	1801	CTATTTCAGAAATGAGAAAATCTTAACCCAAATAGAAATACTCGAAAAGAAATCTGCTTTCA	1860
OY	1861	AAACGAAAGCTGAACCTATAAAGCAGCATATAGCAATATGGAACCTGCAATTAATATATCC	1920
Db	1861	AAACGAAAGCTGAACCTATAAAGCAGCATATAGCAATATGGAACCTGCAATTAATATATCC	1920
OY	1921	ACAAATTCAAAAGCACCTTAAAGAAATAGGCGTGAAGAGAAAGTCTTACACGAGCATATTC	1980
Db	1921	ACAAATTCAAAAGCACCTTAAAGAAATAGGCGTGAAGAGAAAGTCTTACACGAGCATATTC	1980
OY	1981	ATGCGCTTGAACCTAGTAGTACGTAGAAATCTAAGCCACCTAATTTGTAAGTAATGGCAAA	2040
Db	1981	ATGCGCTTGAACCTAGTAGTACGTAGAAATCTAAGCCACCTAATTTGTAAGTAATGGCAAA	2040
OY	2041	TTGATATGTTGTTACGAGTAGGAAGATTAAGAAAAGAAAAAGTACAAACCAATGGCAGATCA	2100
Db	2041	TTGATATGTTGTTACGAGTAGGAAGATTAAGAAAAGAAAAAGTACAAACCAATGGCAGATCA	2100
OY	2101	GGCAGACAGAAAACCTACAACTCATGGAAGTAAAGAACCTGCACATCTGGAGCCAAAGAA	2160
Db	2101	GGCAGACAGAAAACCTACAACTCATGGAAGTAAAGAACCTGCACATCTGGAGCCAAAGAA	2160
OY	2161	GTAAACAAGCCAAATGAAACAGACAAGTAAGAAACATGACAGTAACTTTCCCAAGGTGA	2220
Db	2161	GTAAACAAGCCAAATGAAACAGACAAGTAAGAAACATGACAGTAACTTTCCCAAGGTGA	2220
OY	2221	AGTTAAACAATGACACCTGCTTTCTTTACTAAGTCTTCAAAATACCAAGTAAAGAAAT	2280
Db	2221	AGTTAAACAATGACACCTGCTTTCTTTACTAAGTCTTCAAAATACCAAGTAAAGAAAT	2280
OY	2281	TTTGTCATCTAGCCTTCCAAAGAGAGAAAAGAGAAACTATGAAACAGTTAAAGTGT	2340
Db	2281	TTTGTCATCTAGCCTTCCAAAGAGAGAAAAGAGAAACTATGAAACAGTTAAAGTGT	2340

QY	2341	CTAATATATGCTGGAGAGACCCCAAAGATCTCATGTTTAAGTGGAGAAAGGGTTTGGCAACTG	2400
Db	2341	CTAAATAATGCTGGAAGACCCCAAAGATCTCATGTTTAAGTGGAGAAAGGGTTTGGCAACTG	2400
QY	2401	AAAGTCTGTGAGAGTGTGACAGTATTTTCATTTGTGTAACGGAGTACGATTATGAGCTCAGG	2460
Db	2401	AAAGATCTGTGAGAGTGTGACAGTATTTTCATTTGTGTAACGGAGTACGATTATGAGCTCAGG	2460
QY	2461	AAAGTATCTGTTACTGGAGTGTAGACCTCTAGGGAAGGCAAAAAACAGAACCAATTAAT	2520
Db	2461	AAAGTATCTGTTACTGGAGTGTAGACCTCTAGGGAAGGCAAAAAACAGAACCAATTAAT	2520
QY	2521	GTGTGAGTCACTGTGACAGCATTTGAAAAACCCCAAGGCACTAATTCATGTTGTTCCAAAG	2580
Db	2521	GTGTGAGTCACTGTGACAGCATTTGAAAAACCCCAAGGCACTAATTCATGTTGTTCCAAAG	2580
QY	2581	ATATATAGAAATGACACAGAGGCTTTAATGATTCATTGGGACATGAAGTTAACCACAGTC	2640
Db	2581	ATATATAGAAATGACACAGAGGCTTTAATGATTCATTGGGACATGAAGTTAACCACAGTC	2640
QY	2641	GGGAACACAGATATGAANAATGGAGAAAGATGTAATGTGTCAGTATTTGGCAGAAATCAT	2700
Db	2641	GGGAACACAGATATGAANAATGGAGAAAGATGTAATGTGTCAGTATTTGGCAGAAATCAT	2700
QY	2701	TCAGGTTTCAAAGGCCAGTCATTTGCTGTGTTTCAAATCCAGAAATGCAGAGAAGG	2760
Db	2701	TCAGGTTTCAAAGGCCAGTCATTTGCTGTGTTTCAAATCCAGAAATGCAGAGAAGG	2760
QY	2761	AATGTGCAACATTTCTCTGCCCACTGTGGGTCTTAAAGAAACAAAGTCCAAACTCACTT	2820
Db	2761	AATGTGCAACATTTCTCTGCCCACTGTGGGTCTTAAAGAAACAAAGTCCAAACTCACTT	2820
QY	2821	TTGATGTGACAACAAAGAGAAAGAAATCAAGSAAAGATGACTTAATATCAAGCTCTAC	2880
Db	2821	TTGATGTGACAACAAAGAGAAAGAAATCAAGSAAAGATGACTTAATATCAAGCTCTAC	2880
QY	2881	AGACAGTTAATATCACTGACAGGCTTTCCTGTGGTTGGTCAGAAAGATTAAGCCAGTTGATA	2940
Db	2881	AGACAGTTAATATCACTGACAGGCTTTCCTGTGGTTGGTCAGAAAGATTAAGCCAGTTGATA	2940
QY	2941	ATGCCAATGTAGTATCAAAAGSAGCTCTAGGTTTGTCTATCATCTTCAGTTCCAGAGCA	3000
Db	2941	ATGCCAATGTAGTATCAAAAGSAGCTCTAGGTTTGTCTATCATCTTCAGTTCCAGAGCA	3000
QY	3001	ACGAACCTGGAGCTATTCTCCAAATBAACATGAGCTTTTCAAAACCCCATATGCTATAC	3060
Db	3001	ACGAACCTGGAGCTATTCTCCAAATBAACATGAGCTTTTCAAAACCCCATATGCTATAC	3060
QY	3061	CACCACTTTTCCCATCAAGTCAATTTGTTAAACATAATGTAAAGAAAAATCTGCTAGAGG	3120
Db	3061	CACCACTTTTCCCATCAAGTCAATTTGTTAAACATAATGTAAAGAAAAATCTGCTAGAGG	3120
QY	3121	AAAACCTTGAAGAACATTCATATGTCACCTGAAAGAAATGGGAATGAGAACCTTCCAA	3180
Db	3121	AAAACCTTGAAGAACATTCATATGTCACCTGAAAGAAATGGGAATGAGAACCTTCCAA	3180
QY	3181	GTACAGTGAACAATTTGACCGTATATACTTAGAGAAAATGTTTAAAGAACCCAGCT	3240
Db	3181	GTACAGTGAACAATTTGACCGTATATACTTAGAGAAAATGTTTAAAGAACCCAGCT	3240
QY	3241	CAAGCAATATTAATGAAGTAGTTCACAGTACTAATGAAGTGGGCTCCAGTATTAATGAA	3300
Db	3241	CAAGCAATATTAATGAAGTAGTTCACAGTACTAATGAAGTGGGCTCCAGTATTAATGAA	3300
QY	3301	TAGGTTCCAGTATGAAGAAACATTCACAGAACATGAGTATGAAACAGAGGCCCCAAATTTGA	3360
Db	3301	TAGGTTCCAGTATGAAGAAACATTCACAGAACATGAGTATGAAACAGAGGCCCCAAATTTGA	3360
QY	3361	ATGCTATGCTTATGATTTAGGGGTTTTCGACCTGAGGCTCTATTAACCAAAAGTCTTCCTGGAA	3420
Db	3361	ATGCTATGCTTATGATTTAGGGGTTTTCGACCTGAGGCTCTATTAACCAAAAGTCTTCCTGGAA	3420
QY	3421	GTAATTTGATGACATCTCGAATAATAAAAAAGCAAGAAATATGAAGAAGTAGTTCAGACTGTTA	3480

```
Db 3421 GTAAATGTAAGCATCCGGAATATAAAGAGAAATGAAAGAGTGGTTCAGACTCTTA 3480
OY 3481 ATACAGATTTCTCTCCATATCTGATTTACAGATACCTAGAAAGCCATGGAAGTATGTC 3540
Db 3481 ATACAGATTTCTCTCCATATCTGATTTACAGATACCTAGAAAGCCATGGAAGTATGTC 3540
OY 3541 ATGCATCTCAGGTTTGTCTGAGACACCTGATGACCTGTTAGATGATGTAATAAAG 3600
Db 3541 ATGCATCTCAGGTTTGTCTGAGACACCTGATGACCTGTTAGATGATGTAATAAAG 3600
OY 3601 AAGATCTAGTGTGTAAGAAATGACATTAAGAAAGTCTGCTGTTTTAGCAAAAGCG 3660
Db 3601 AAGATCTAGTGTGTAAGAAATGACATTAAGAAAGTCTGCTGTTTTAGCAAAAGCG 3660
OY 3661 TCCAGAAAGAGAGAGCTTACAGAGAGTCTAGCCCTTACACCATACATTTGGCTCAG 3720
Db 3661 TCCAGAGAGAGAGCTTACAGAGAGTCTAGCCCTTACACCATACATTTGGCTCAG 3720
OY 3721 GTTACCGAAGAGAGGCGCAAGAAATTAAGATCTCTCAGAGAGAACTTATCTAGAGATG 3780
Db 3721 GTTACCGAAGAGAGGCGCAAGAAATTAAGATCTCTCAGAGAGAACTTATCTAGAGATG 3780
OY 3781 AAGAGCTTCCCTGCTTCCAACTTGTATTTGTAAGTAACAAATATACCTTCAGT 3840
Db 3781 AAGAGCTTCCCTGCTTCCAACTTGTATTTGTAAGTAACAAATATACCTTCAGT 3840
OY 3841 CTACTAGGCAATAGCACCGTTGCTACCGAGTGTCTCTAAGAACACAGAGAGAAATTTAT 3900
Db 3841 CTACTAGGCAATAGCACCGTTGCTACCGAGTGTCTCTAAGAACACAGAGAGAAATTTAT 3900
OY 3901 TATCATTTGAAGATAGCTTAATGACTGCAAGTACCAAGGTAATATTGGCAAGGCAATCTC 3960
Db 3901 TATCATTTGAAGATAGCTTAATGACTGCAAGTACCAAGGTAATATTGGCAAGGCAATCTC 3960
OY 3961 AGCAACATACCTTGTAGAGAAACAAAAATGTTCTGCTAGCTGTTTCTTCACATGCA 4020
Db 3961 AGCAACATACCTTGTAGAGAAACAAAAATGTTCTGCTAGCTGTTTCTTCACATGCA 4020
OY 4021 GTGAATTTGGAAAGCTTGACATGCAAAATCAAAACACCAGATCCCTTCTTGATTTGTTCTT 4080
Db 4021 GTGAATTTGGAAAGCTTGACATGCAAAATCAAAACACCAGATCCCTTCTTGATTTGTTCTT 4080
OY 4081 CCAACCAATGAGGATCATGCTGTAAGAGCCAGGAGTGGTCTGAGTGACAGAGAAATGG 4140
Db 4081 CCAACCAATGAGGATCATGCTGTAAGAGCCAGGAGTGGTCTGAGTGACAGAGAAATGG 4140
OY 4141 TTTTCAGATGATGAAGAAAGAGAGAGCGGCTTGGAAAGAAATATCAAGAAAGCAAGCA 4200
Db 4141 TTTTCAGATGATGAAGAAAGAGAGAGCGGCTTGGAAAGAAATATCAAGAAAGCAAGCA 4200
OY 4201 TGGATTTCAAACTTTAGGTGAAGAGCATCTGGGTGTGAGAGTAAACAAACGCTCTCGAAG 4260
Db 4201 TGGATTTCAAACTTTAGGTGAAGAGCATCTGGGTGTGAGAGTAAACAAACGCTCTCGAAG 4260
OY 4261 ACTGCTCAGGCGTATCTCTCTCAGAGTGCATTTTAAACACTCAGCAGAGAGATACCATGC 4320
Db 4261 ACTGCTCAGGCGTATCTCTCTCAGAGTGCATTTTAAACACTCAGCAGAGAGATACCATGC 4320
OY 4321 AACATTAACCTGATTAAGGTCGAGAGAAATGGCGAATAGAACTGTGTTAGAAAGCAG 4380
Db 4321 AACATTAACCTGATTAAGGTCGAGAGAAATGGCGAATAGAACTGTGTTAGAAAGCAG 4380
OY 4381 ATGGAGAGCAGCCTTTTAACAGCTACCTTCCATATAAGTACTCTTCTGCTTGAAG 4440
Db 4381 ATGGAGAGCAGCCTTTTAACAGCTACCTTCCATATAAGTACTCTTCTGCTTGAAG 4440
OY 4441 ACCGCGCAATTCAGAAACAAAGACATCAGAAAAGAGTATTACTTCACAGAAAAGTA 4500
Db 4441 ACCGCGCAATTCAGAAACAAAGACATCAGAAAAGAGTATTACTTCACAGAAAAGTA 4500
OY 4501 GTGAATGCCCTATTAAGCCAGAAATCCAGAAAGCCTTTCTGCTGACAAAGTTTGAGGTCTG 4560
Db 4501 GTGAATGCCCTATTAAGCCAGAAATCCAGAAAGCCTTTCTGCTGACAAAGTTTGAGGTCTG 4560
```



OY 5641 GTGAGCACTGTACCAAGTGCAGAGCGTGACACTTAATCCCATGCCAGATCCCCACA 5700  
|||  
Db 5641 GTGAGCACTGTACCAAGTGCAGAGCGTGACACTTAATCCCATGCCAGATCCCCACA 5700  
5641 GTGAGCACTGTACCAAGTGCAGAGCGTGACACTTAATCCCATGCCAGATCCCCACA 5700

OY 5701 GCCACTACTGA 5711  
|||||  
Db 5701 GCCACTACTGA 5711

## RESULT 6

US-09-734-672-1  
: Sequence 1, Application US/09734672

Publication No. US20020183268A1  
GENERAL INFORMATION:

TITLE OF INVENTION: Coding Sequences of the  
 BRCA1 Gene  
 NUMBER OF SEQUENCES: 72  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Morgan Lewis & Bockius LLP  
 STREET: 1111 Pennsylvania Ave., N.W.  
 CITY: Washington  
 STATE: District of Columbia  
 COUNTRY: USA  
 ZIP: 20004

SEQUENCE DESCRIPTION:	SEQ ID NO:	1:
US-09-734-672-1		

QY 1 AGCTCGTGAGACTTCCTGGACCCCGCACCGCTGTGGGTTTCTCAGATAACTGGCC 60

Db	1	AGCTGGCTGAGACCTTCCTGGACCCCGACACAGGGCTGTGGGGTTTCTCAGATTAACGTGGGCG	60
OY	61	CCTGGCGCTCAGAGAGGCTTTCACCCCTCTGCTGGGTAAAGTTCAATTTGGAACAGAAAGAA	120
Db	61	CCTGGCGCTCAGAGAGGCGCTTTCACCCCTCTGCTGGGTAAAGTTCAATTTGGAACAGAAAGAA	120
OY	121	TGGATTTTATCTGCTCTTGGCGTTGGAAGAAATACAATAATGTCAATTAATCTATATCGAGAAA	180
Db	121	TGGATTTTATCTGCTCTTGGCGTTGGAAGAAATACAATAATGTCAATTAATCTATATCGAGAAA	180
OY	181	TCTTAGAGTGTCCATCTGTCTGGAGTTGATCAAGAAACGTGTCTCCACAAAGTGTGACC	240
Db	181	TCTTAGAGTGTCCATCTGTCTGGAGTTGATCAAGAAACGTGTCTCCACAAAGTGTGACC	240
OY	241	ACATATTTTGGAAATTTTGCAATGCTGGAACCTTCCAACDACAAGAAAGGGCGTTACAGT	300
Db	241	ACATATTTTGGAAATTTTGGCAATGCTGGAACCTTCCAACDACAAGAAAGGGCGTTACAGT	300
OY	301	GTCCCTTTATGTAAGATATGATTAACCAAAAGAGCGCTCAAGAAAGTACGAGATTTAGTC	360
Db	301	GTCCCTTTATGTAAGATATGATTAACCAAAAGAGCGCTCAAGAAAGTACGAGATTTAGTC	360
OY	361	AACTTGTTGAAGAGCTATTTGAATAATCATTTTGCTTTTCACCTTGACACAGGTTTGGAGT	420
Db	361	AACTTGTTGAAGAGCTATTTGAATAATCATTTTGCTTTTCACCTTGACACAGGTTTGGAGT	420
OY	421	ATGCAAAACAGCTATTAATTTTGCACAAAAGGAAATAACTGTCGTAACATGCTAAAGATG	480
Db	421	ATGCAAAACAGCTATTAATTTTGCACAAAAGGAAATAACTGTCGTAACATGCTAAAGATG	480
OY	481	AAGTTTCTATCATGCCAAAGATATGGGCTACAGAAACCGTGCCAAAAGACTTTACAGAGTG	540
Db	481	AAGTTTCTATCATGCCAAAGATATGGGCTACAGAAACCGTGCCAAAAGACTTTACAGAGTG	540
OY	541	AAACCCGAAATATCTCTTCTTTCAGAGAAACCAAGTCTAGTGTCCAACTCTCTAACTTTGGA	600
Db	541	AAACCCGAAATATCTCTTCTTTCAGAGAAACCAAGTCTAGTGTCCAACTCTCTAACTTTGGA	600
OY	601	CTGTGAGAACTCTGAGACGACAAAGAGCGCGATACAACTCTCAACCTCTCTCAATTG	660
Db	601	CTGTGAGAACTCTGAGACGACAAAGAGCGCGATACAACTCTCAACCTCTCTCAATTG	660
OY	661	AATTGGGATCTGATTTCTTCTGAAGATACCGTTAATAAAGCAACTTAATGCAAGTGTGGAG	720
Db	661	AATTGGGATCTGATTTCTTCTGAAGATACCGTTAATAAAGCAACTTAATGCAAGTGTGGAG	720
OY	721	ATCAAGAAATTTTTCAAATCAACCCCTCAAGAACACGAGGATGAATCATGTTGGATTCTG	780
Db	721	ATCAAGAAATTTTTCAAATCAACCCCTCAAGAACACGAGGATGAATCATGTTGGATTCTG	780
OY	781	CAAAAAGAGCTGTGTGAAATTTTCTGAGACGAGATTAACAATACTGAACATCAAC	840
Db	781	CAAAAAGAGCTGTGTGAAATTTTCTGAGACGAGATTAACAATACTGAACATCAAC	840
OY	841	CCAGTAATATGATTTGAAACACCACTGAAAGCGTGAAGCGATCCAGAAAAGT	900
Db	841	CCAGTAATATGATTTGAAACACCACTGAAAGCGTGAAGCGATCCAGAAAAGT	900
OY	901	ATCAGGGATGTTGTTTCAAACTTGCATGTGGACCAATGTGGCAAAATCTCATGGCA	960
Db	901	ATCAGGGATGTTGTTTCAAACTTGCATGTGGACCAATGTGGCAAAATCTCATGGCA	960
OY	961	GCTCATTAACAGCATGAGAACAGCAAGTTATTACTCACTAAAGACGAATGTAGTGA	1020
Db	961	GCTCATTAACAGCATGAGAACAGCAAGTTATTACTCACTAAAGACGAATGTAGTGA	1020
OY	1021	AGGCTGAATTTGTGAATTAAGCAAAACGCTGGCTTGAAGAGGCCAACATAACGAT	1080
Db	1021	AGGCTGAATTTGTGAATTAAGCAAAACGCTGGCTTGAAGAGGCCAACATAACGAT	1080
OY	1081	GGGCTGGAAGTAAAGAAAACATGTATGATTAAGCGGACTCCAGCACAGAAAAAAGTATG	1140
Db	1081	GGGCTGGAAGTAAAGAAAACATGTATGATTAAGCGGACTCCAGCACAGAAAAAAGTATG	1140



Db 1081 GGGCTGGAAGTAMGAAACAATGATATAGCGGACTCCCGACGACAGAAAAAGGTAG 1140  
Qy 1141 ATCTGAATGCGATCCCTGTGTGAGAAAAAGATGGATAGCAGAACTGCGATGCT 1200  
Db 1141 ATCTGAATGCGATCCCTGTGTGAGAAAAAGATGGATAGCAGAACTGCGATGCT 1200  
Qy 1201 CAGAGAACTCTAGAGATCTGAAAGATGTTCTCTGGATTAACACTAAATAGCAGATTGCA 1260  
Db 1201 CAGAGAACTCTAGAGATCTGAAAGATGTTCTCTGGATTAACACTAAATAGCAGATTGCA 1260  
Qy 1261 AAGTATATGAGTGTGTTTCCAGAAAGTATGATGACTGTTAGTTCTGATGACTCAGATGATG 1320  
Db 1261 AAGTATATGAGTGTGTTTCCAGAAAGTATGATGACTGTTAGTTCTGATGACTCAGATGATG 1320  
Qy 1321 GGGAGTGTGATCAATGCGAAAGTATGATGATGATGATGATGATGATGATGATGATGATG 1380  
Db 1321 GGGAGTGTGATCAATGCGAAAGTATGATGATGATGATGATGATGATGATGATGATGATG 1380  
Qy 1381 AATATCTGCTTCTTCAGAGAAATATGATGATGATGATGATGATGATGATGATGATGATG 1440  
Db 1381 AATATCTGCTTCTTCAGAGAAATATGATGATGATGATGATGATGATGATGATGATGATG 1440  
Qy 1441 TATGTAAAGTGAAGAGTCTCCTCAATCAGTAGAGATTAATATTGAGACAAATAT 1500  
Db 1441 TATGTAAAGTGAAGAGTCTCCTCAATCAGTAGAGATTAATATTGAGACAAATAT 1500  
Qy 1501 TTGGGAAACCTATCGGAAGAGGACGCTCCCACTTAAGGCTATGATGATGATGATGATG 1560  
Db 1501 TTGGGAAACCTATCGGAAGAGGACGCTCCCACTTAAGGCTATGATGATGATGATGATG 1560  
Qy 1561 TAAATATAGAGCATTTTGTACTGAGCCAGACAGATTAATACAGAGCGTCCCTCAATA 1620  
Db 1561 TAAATATAGAGCATTTTGTACTGAGCCAGACAGATTAATACAGAGCGTCCCTCAATA 1620  
Qy 1621 AATTAAGGCTAAAGAGAGTCTATGAGGCTTCACTGAGATTTTATCAAGAAAG 1680  
Db 1621 AATTAAGGCTAAAGAGAGTCTATGAGGCTTCACTGAGATTTTATCAAGAAAG 1680  
Qy 1681 CAGATTTGGCAGTTCAAAAGACTCTGAAATGATTAATCAAGAGACTAACCAGAGAGC 1740  
Db 1681 CAGATTTGGCAGTTCAAAAGACTCTGAAATGATTAATCAAGAGACTAACCAGAGAGC 1740  
Qy 1741 AGATGCTCAAGTATGATTAATTAATAGTGTGATGAGATTAACAAAGAGTAT 1800  
Db 1741 AGATGCTCAAGTATGATTAATTAATAGTGTGATGAGATTAACAAAGAGTAT 1800  
Qy 1801 CTATTCAGATGAGAAATATCTTAACCCATAGAAATGATGATGATGATGATGATGATG 1860  
Db 1801 CTATTCAGATGAGAAATATCTTAACCCATAGAAATGATGATGATGATGATGATGATG 1860  
Qy 1861 AAAGCAAGGCTGAACCTATTAAGCAGATTAATGATGATGATGATGATGATGATGATG 1920  
Db 1861 AAAGCAAGGCTGAACCTATTAAGCAGATTAATGATGATGATGATGATGATGATGATG 1920  
Qy 1921 ACAATTCAAAAGCACTTAAGAAATAGGCTGAGAGAGAGTCTTCTACAGAGCATATTC 1980  
Db 1921 ACAATTCAAAAGCACTTAAGAAATAGGCTGAGAGAGAGTCTTCTACAGAGCATATTC 1980  
Qy 1981 ATGCGCTGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2040  
Db 1981 ATGCGCTGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2040  
Qy 2041 TTGATAGTGTCTAGCAGTGAAGAGATTAAGAAAAAGTACAACTGAGGAGCAATGCA 2100  
Db 2041 TTGATAGTGTCTAGCAGTGAAGAGATTAAGAAAAAGTACAACTGAGGAGCAATGCA 2100  
Qy 2101 GGCACAGCAGAAACCTTAACATCATGAGAGGTAAGAACTGCACTGAGGAGCAAGAGA 2160  
Db 2101 GGCACAGCAGAAACCTTAACATCATGAGAGGTAAGAACTGCACTGAGGAGCAAGAGA 2160  
Qy 2161 GTRACAGGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220  
Db 2161 GTRACAGGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220

Qy 2221 AGTTAACAAATGACACCTGTTCTTTTACTAGTGTCAATATACAGTGAATTAAGAT 2280  
Db 2221 AGTTAACAAATGACACCTGTTCTTTTACTAGTGTCAATATACAGTGAATTAAGAT 2280  
Qy 2281 TTGTCAATCTAGCCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2340  
Db 2281 TTGTCAATCTAGCCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2340  
Qy 2341 CTATATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2400  
Db 2341 CTATATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2400  
Qy 2401 AAAGATCTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2460  
Db 2401 AAAGATCTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2460  
Qy 2461 AAAGATCTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520  
Db 2461 AAAGATCTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520  
Qy 2521 GTGTGAGTCAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580  
Db 2521 GTGTGAGTCAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580  
Qy 2581 ATATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2640  
Db 2581 ATATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2640  
Qy 2641 GGGAAACAGCATAGAAATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2700  
Db 2641 GGGAAACAGCATAGAAATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2700  
Qy 2701 TCAAGGTTTCAAGGCGCATCTTGTCTGTCTTCAATCAGAGAGAGAGAGAGAGAG 2760  
Db 2701 TCAAGGTTTCAAGGCGCATCTTGTCTGTCTTCAATCAGAGAGAGAGAGAGAGAG 2760  
Qy 2761 AATGTGCAACATCTCTGCGCAGCTGAGGCTTCTTCAAGAGAGAGAGAGAGAGAG 2820  
Db 2761 AATGTGCAACATCTCTGCGCAGCTGAGGCTTCTTCAAGAGAGAGAGAGAGAGAG 2820  
Qy 2821 TTGATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2880  
Db 2821 TTGATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2880  
Qy 2881 AGACAGTATATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2940  
Db 2881 AGACAGTATATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2940  
Qy 2941 ATGCAAAATGATATCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3000  
Db 2941 ATGCAAAATGATATCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3000  
Qy 3001 ACAGAACTGAGCATATTAACCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060  
Db 3001 ACAGAACTGAGCATATTAACCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060  
Qy 3061 CACCACTTTTCCATCAAGTCAATTTGTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 3120  
Db 3061 CACCACTTTTCCATCAAGTCAATTTGTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 3120  
Qy 3121 AAACCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3180  
Db 3121 AAACCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3180  
Qy 3181 GTACAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3240  
Db 3181 GTACAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3240  
Qy 3241 CAAGCAATATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3300  
Db 3241 CAAGCAATATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3300

OY 3301 TAGTCCAGTATGAAAAACATTCAGACGAACTAGTAGAACAAGAGGCCCAAAATTGA 3360  
 |||||  
 Db 3301 TAGTCCAGTATGAAAAACATTCAGACGAACTAGTAGAACAAGAGGCCCAAAATTGA 3360  
 OY 3361 ATGTCATGCTTAAGATTAAGGGTTTTCGAACCTGAGGCTATTAACAAGTCTTCTGGAA 3420  
 |||||  
 Db 3361 ATGTCATGCTTAAGATTAAGGGTTTTCGAACCTGAGGCTATTAACAAGTCTTCTGGAA 3420  
 OY 3421 GTAATTGTAACATCTGTAATTAATAAAGCAAGATATGAAGAGTACTGACCTGTA 3480  
 |||||  
 Db 3421 GTAATTGTAACATCTGTAATTAATAAAGCAAGATATGAAGAGTACTGACCTGTA 3480  
 OY 3481 ATACAGATTTCTCTCCATATCTGATTTCAAGATTAAGAACGCTTAAGGAACTACTC 3540  
 |||||  
 Db 3481 ATACAGATTTCTCTCCATATCTGATTTCAAGATTAAGAACGCTTAAGGAACTACTC 3540  
 OY 3541 ATGATCTCAGGTTTGTCTGACACCTGATGACCTGTTGATGATGATGATGATGATGATG 3600  
 |||||  
 Db 3541 ATGATCTCAGGTTTGTCTGACACCTGATGACCTGTTGATGATGATGATGATGATGATG 3600  
 OY 3601 AAGATCTAGTCTGTAATTAAGAAAGTCTGCTGTTTTCAGCAAAAGCG 3660  
 |||||  
 Db 3601 AAGATCTAGTCTGTAATTAAGAAAGTCTGCTGTTTTCAGCAAAAGCG 3660  
 OY 3661 TCCAGAAAGAGAGCTTACGAGAGTCCCTAGCCCTTCACCCATACACATTTGGCTCAG 3720  
 |||||  
 Db 3661 TCCAGAAAGAGAGCTTACGAGAGTCCCTAGCCCTTCACCCATACACATTTGGCTCAG 3720  
 OY 3721 GTTACCGAAGAGGGGCCAAGAAATTAAGTCTCAGAAAGAACTTATCTAGTGAAGATG 3780  
 |||||  
 Db 3721 GTTACCGAAGAGGGGCCAAGAAATTAAGTCTCAGAAAGAACTTATCTAGTGAAGATG 3780  
 OY 3781 AAGAGCTCCCTGCTCCAACTGTTATTTGGTAAAGTAAACAATATACCTCTCAGT 3840  
 |||||  
 Db 3781 AAGAGCTCCCTGCTCCAACTGTTATTTGGTAAAGTAAACAATATACCTCTCAGT 3840  
 OY 3841 CTACAGGATAGACACCGTCTACCGAGTCTGTCTAAGAACACAGAGAGAGATTTAT 3900  
 |||||  
 Db 3841 CTACAGGATAGACACCGTCTACCGAGTCTGTCTAAGAACACAGAGAGAGATTTAT 3900  
 OY 3901 TATGATTAAGAAATAGCTTAATGACATGACATGACATGACATGACATGACATGACATG 3960  
 |||||  
 Db 3901 TATGATTAAGAAATAGCTTAATGACATGACATGACATGACATGACATGACATGACATG 3960  
 OY 3961 AGGAACATCACCCTTAGTAGAGAAACAAATATCTGCTAGCTGTTCTTCTCAAGTGA 4020  
 |||||  
 Db 3961 AGGAACATCACCCTTAGTAGAGAAACAAATATCTGCTAGCTGTTCTTCTCAAGTGA 4020  
 OY 4021 GTGATTTGGAAGACTTACCTGCAAAATACAAACACCCAGAGATCTTCTTGAATGGTCTT 4080  
 |||||  
 Db 4021 GTGATTTGGAAGACTTACCTGCAAAATACAAACACCCAGAGATCTTCTTGAATGGTCTT 4080  
 OY 4081 CCAAAACAATGAGGATCAGTCTGAAGCCGAGGAGTGTCTGAGTACCAAGAAATGG 4140  
 |||||  
 Db 4081 CCAAAACAATGAGGATCAGTCTGAAGCCGAGGAGTGTCTGAGTACCAAGAAATGG 4140  
 OY 4141 TTTTCAGATGATGAAGAAAGAGAACGGCTTGAAGAAATATCAAGAAAGAGCA 4200  
 |||||  
 Db 4141 TTTTCAGATGATGAAGAAAGAGAACGGCTTGAAGAAATATCAAGAAAGAGCA 4200  
 OY 4201 TGGATTAACCTTAGTGAAGACAGCATCTGGGTGTGAAGTGAAGAAACAGCTCTGAAG 4260  
 |||||  
 Db 4201 TGGATTAACCTTAGTGAAGACAGCATCTGGGTGTGAAGTGAAGAAACAGCTCTGAAG 4260  
 OY 4261 ACTGCTCAGGCTATCTCTCAGAGTGAACATTTTAACCACTCAGCAGAGGATACCATGC 4320  
 |||||  
 Db 4261 ACTGCTCAGGCTATCTCTCAGAGTGAACATTTTAACCACTCAGCAGAGGATACCATGC 4320  
 OY 4321 AACATAACCTGATTAAGCTCAGCAGAAATGCTGTAAGTGAAGTGTGTTAAGACAGC 4380  
 |||||  
 Db 4321 AACATAACCTGATTAAGCTCAGCAGAAATGCTGTAAGTGAAGTGTGTTAAGACAGC 4380  
 OY 4381 ATGGAGCCAGCCTTTCAACGACTACCTTCCATCATTAAGTGAATGACTTCTGACCCTTGA 4440  
 |||||

Db 4381 ATGGAGCCACACTTCTTAACGACTACCTTCATCATTAAGTGAAGTCTCTGACCTTGAGG 4440  
 |||||  
 OY 4441 ACCTGCCGAATCCGAACAAAGACATCAAGAAAAAGAGTAACTTACCTCAGCAAAACTA 4500  
 |||||  
 Db 4441 ACCTGCCGAATCCGAACAAAGACATCAAGAAAAAGAGTAACTTACCTCAGCAAAACTA 4500  
 OY 4501 GTGAATACCTTAATGAAGCAGATCCAGAGGCTTCTGCTGACAAAGTTGAGGCTGTG 4560  
 |||||  
 Db 4501 GTGAATACCTTAATGAAGCAGATCCAGAGGCTTCTGCTGACAAAGTTGAGGCTGTG 4560  
 OY 4561 CAGATAGTTCTACAGATTAATAAAGAACAGAGAGTGAAGAGTCAATCCCTCTTAAT 4620  
 |||||  
 Db 4561 CAGATAGTTCTACAGATTAATAAAGAACAGAGAGTGAAGAGTCAATCCCTCTTAAT 4620  
 OY 4621 GCGATAGTTCTACAGATTAATAAAGAACAGAGAGTGAAGAGTCAATCCCTCTTAAT 4680  
 |||||  
 Db 4621 GCGATAGTTCTACAGATTAATAAAGAACAGAGAGTGAAGAGTCAATCCCTCTTAAT 4680  
 OY 4681 GCGATAGTTCTACAGATTAATAAAGAACAGAGAGTGAAGAGTCAATCCCTCTTAAT 4740  
 |||||  
 Db 4681 GCGATAGTTCTACAGATTAATAAAGAACAGAGAGTGAAGAGTCAATCCCTCTTAAT 4740  
 OY 4741 AGTCTGGGCTCAGCAGATTTGACGGAACATCTTACTGCAAGGCAAGATCTAGAGGAA 4800  
 |||||  
 Db 4741 AGTCTGGGCTCAGCAGATTTGACGGAACATCTTACTGCAAGGCAAGATCTAGAGGAA 4800  
 OY 4801 CCCCTTACCTGGAATCTGGAATCAGCTCTTCTGATGACCTCTGAATCTATCTCTG 4860  
 |||||  
 Db 4801 CCCCTTACCTGGAATCTGGAATCAGCTCTTCTGATGACCTCTGAATCTATCTCTG 4860  
 OY 4861 AAGACAGAGCCCGAGAGTCACTGCTGTGGCAACATCACTTCAACCTCTGATGA 4920  
 |||||  
 Db 4861 AAGACAGAGCCCGAGAGTCACTGCTGTGGCAACATCACTTCAACCTCTGATGA 4920  
 OY 4921 AAGTCCCAATTTGAAGTCCAGAAATCTGCCAGAGTCCAGAGTCCATCTACTACTG 4980  
 |||||  
 Db 4921 AAGTCCCAATTTGAAGTCCAGAAATCTGCCAGAGTCCAGAGTCCATCTACTACTG 4980  
 OY 4981 ATACTGCTGGGTATTAATGCAATGGAAGAACTGTGAGAGAGGAGAGGCAAAATTTGAC 5040  
 |||||  
 Db 4981 ATACTGCTGGGTATTAATGCAATGGAAGAACTGTGAGAGAGGAGAGGCAAAATTTGAC 5040  
 OY 5041 CTTCAACGAAAGGTCMAACAAAGAAATGTCATGCTGCTGAGCCCTGAGCCAGAG 5100  
 |||||  
 Db 5041 CTTCAACGAAAGGTCMAACAAAGAAATGTCATGCTGCTGAGCCCTGAGCCAGAG 5100  
 OY 5101 AATTATGCTGCTGACAAATTTGCGCAGAAACACCATCACTTAATCTAATTA 5160  
 |||||  
 Db 5101 AATTATGCTGCTGACAAATTTGCGCAGAAACACCATCACTTAATCTAATTA 5160  
 OY 5161 CTGAAGAGACTACATCTGTTATGAAGAAAGATGCTGATGCTGCTGCTGCTGCTGCTG 5220  
 |||||  
 Db 5161 CTGAAGAGACTACATCTGTTATGAAGAAAGATGCTGATGCTGCTGCTGCTGCTGCTG 5220  
 OY 5221 TGAATATTTCTGAGAAATTCGGGAGAGAAATGAGTATGATGCTGCTGCTGCTGCTG 5280  
 |||||  
 Db 5221 TGAATATTTCTGAGAAATTCGGGAGAGAAATGAGTATGATGCTGCTGCTGCTGCTG 5280  
 OY 5281 AGTCTATTTAAAGAAAGAAATGCTGAATGAGCATGATTTGAAGTCAAGAGATGAG 5340  
 |||||  
 Db 5281 AGTCTATTTAAAGAAAGAAATGCTGAATGAGCATGATTTGAAGTCAAGAGATGAG 5340  
 OY 5341 TCAATGGAAGAAACCCAGAGGTCGAAAGCGAGAGAGAAATCCAGAGAGAAAGATCT 5400  
 |||||  
 Db 5341 TCAATGGAAGAAACCCAGAGGTCGAAAGCGAGAGAGAAATCCAGAGAGAAAGATCT 5400  
 OY 5401 TCAGGGGCTGGAATCTGTTGCTATGAGGCTTCAACCAATCTCCACAGATCAACTGG 5460  
 |||||  
 Db 5401 TCAGGGGCTGGAATCTGTTGCTATGAGGCTTCAACCAATCTCCACAGATCAACTGG 5460  
 OY 5461 AATGATGATACAGCTGTGTGCTCTCTGCTGTAAGAGAGCTTTCATCTACCTTGG 5520  
 |||||

Db 5461 AATGATGATACAGCTGTGTGCTCTTCTGTGTGAAGAGCTTTCATTCACCTTGG 5520  
 QY 5521 GCACAGGTGTCCACCAATGTGTGTGTGCGACCGACATGCTTGACAGAGACAAATGGCT 5580  
 Db 5521 GCACAGGTGTCCACCAATGTGTGTGTGCGACCGACATGCTTGACAGAGACAAATGGCT 5580  
 QY 5581 TCCATGCAATTTGGCAGATGTGTGAGCACCCTGTGTGACCCGAGAGTGGTGTGGACA 5640  
 Db 5581 TCCATGCAATTTGGCAGATGTGTGAGCACCCTGTGTGACCCGAGAGTGGTGTGGACA 5640  
 QY 5641 GTGTAGCAGCTCTACCACTGCGACAGAGCTGACACCTTACCTGATACCCAGATCCGCCACA 5700  
 Db 5641 GTGTAGCAGCTCTACCACTGCGACAGAGCTGACACCTTACCTGATACCCAGATCCGCCACA 5700  
 QY 5701 GCCACTACTGA 5711  
 Db 5701 GCCACTACTGA.5711

## RESULT 7

US-09-982-828-1

Sequence 1, Application US/09982828

Publication No. US20030022184A1

## GENERAL INFORMATION:

## APPLICANT:

Murphy, Patricia D.  
 Allen, Antonette C.  
 Critz, Brenda S.  
 Olson, Sheri J.  
 Thurber, Denise  
 Zeng, Bin

TITLE OF INVENTION: Coding Sequences of the Human  
 BRCAL Gene

NUMBER OF SEQUENCES: 72  
 CORRESPONDENCE ADDRESS:

ADDRESSEE: Morgan Lewis & Bockius LLP  
 STREET: 1111 Pennsylvania Avenue N. W.  
 CITY: Washington  
 STATE: DC

COUNTRY: USA

ZIP: 20004

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/982.828

FILING DATE: 22-Oct-2001

CLASSIFICATION: &lt;Unknown&gt;

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 09/074.453

FILING DATE: 1998-05-06

APPLICATION NUMBER: US 08/798.691

FILING DATE: 1997-02-12

APPLICATION NUMBER: US 08/598.591

FILING DATE: 1996-02-12

## ATTORNEY/AGENT INFORMATION:

NAME: Michael S. Tuscan

REGISTRATION NUMBER: 43,210

REFERENCE/DOCKET NUMBER: 44921-5053-01-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-739-3000

TELEFAX: 202-739-3001

INFORMATION FOR SEQ ID NO: 1:

## SEQUENCE CHARACTERISTICS:

LENGTH: 5711 base pairs

TYPE: nucleic acid

STRANDEDNESS: not relevant

MOLECULE TYPE: cDNA

ORGANISM: Homo sapiens

STRAIN: BRCAL (om11)  
 POSITION IN GENOME:  
 CHROMOSOME/SEGMENT: 17  
 MAP POSITION: 17q21  
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 us-09-982-828-1

Query Match 99.9% Score 5703; DB 9; Length 5711;  
 Best Local Similarity 99.9% Pred. No. 0;  
 Matches 5706; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AGCTCGCTGAGACTTCTCTGAGACCCGACCGAGCGTGTGGGTTTTCAGATTAACCTGGCC 60  
 Db 1 AGCTCGCTGAGACTTCTCTGAGACCCGACCGAGCGTGTGGGTTTTCAGATTAACCTGGCC 60  
 QY 61 CCTGCGCTCAGAGAGGCTTACACCTCTGCTCTGCGGTAAGTTATTGGAACAGAAAGAA 120  
 Db 61 CCTGCGCTCAGAGAGGCTTACACCTCTGCTCTGCGGTAAGTTATTGGAACAGAAAGAA 120  
 QY 121 TGGATTAATCTGCTCTCGGGTTGAAGAGTACAAATGCAATTAATGCTATGACAGAAA 180  
 Db 121 TGGATTAATCTGCTCTCGGGTTGAAGAGTACAAATGCAATTAATGCTATGACAGAAA 180  
 QY 181 TCTTAGAGTGTCCATCTCTCTGAGAGTTGATCAAGAGAACCTGTCTCCAGAAAGTGTACC 240  
 Db 181 TCTTAGAGTGTCCATCTCTCTGAGAGTTGATCAAGAGAACCTGTCTCCAGAAAGTGTACC 240  
 QY 241 ACATATTTTGAATTTTGTGATGCTGAACCTTCTCAACCAAGAAAGAGGCTTACAGT 300  
 Db 241 ACATATTTTGAATTTTGTGATGCTGAACCTTCTCAACCAAGAAAGAGGCTTACAGT 300  
 QY 301 GTCTTTATGTAAGATGATATTAACCAAGAGAGGCTTCAAGAAATAGAGATTAGTC 360  
 Db 301 GTCTTTATGTAAGATGATATTAACCAAGAGAGGCTTCAAGAAATAGAGATTAGTC 360  
 QY 361 AACTTGTGAAGAGCTATTAATAATGCTTTGACCTTTGACCTGACAGGTTTGGAGT 420  
 Db 361 AACTTGTGAAGAGCTATTAATAATGCTTTGACCTTTGACCTGACAGGTTTGGAGT 420  
 QY 421 ATGCAAAACAGCTATATTTTGAACAAAAGAAATAACTCTCTGAACATCTAAGAGATG 480  
 Db 421 ATGCAAAACAGCTATATTTTGAACAAAAGAAATAACTCTCTGAACATCTAAGAGATG 480  
 QY 481 AAGTTCTATCACCAGATGATGGGCTACGAAACCGTGCAAAAGACTTCTACAGAGT 540  
 Db 481 AAGTTCTATCACCAGATGATGGGCTACGAAACCGTGCAAAAGACTTCTACAGAGT 540  
 QY 541 AACCCGAAATCCTTCCTTCAGAGAAACAGTCTCAAGTCTCAACTCTTAACCTTGGAA 600  
 Db 541 AACCCGAAATCCTTCCTTCAGAGAAACAGTCTCAAGTCTCAACTCTTAACCTTGGAA 600  
 QY 601 CTGTGAGAACTGTGAGAGACAAAGCAGCGATACAACTTAAAAAGAGTGTCTACATTG 660  
 Db 601 CTGTGAGAACTGTGAGAGACAAAGCAGCGATACAACTTAAAAAGAGTGTCTACATTG 660  
 QY 661 AATTGGATCTGATTTCTTGAAGATACCGTTATTAAGCAACTTTTCAGTGGAG 720  
 Db 661 AATTGGATCTGATTTCTTGAAGATACCGTTATTAAGCAACTTTTCAGTGGAG 720  
 QY 721 ATCAGAAATTTGTAACAATCACCCCTCAAGAGAACAGGATGAATCACTTTGGATTGTG 780  
 Db 721 ATCAGAAATTTGTAACAATCACCCCTCAAGAGAACAGGATGAATCACTTTGGATTGTG 780  
 QY 781 CAAAAAAGGCTGCTTGTGAATTTTTCGAGAGCATGTAACTAATCTGAACATCATCAAC 840  
 Db 781 CAAAAAAGGCTGCTTGTGAATTTTTCGAGAGCATGTAACTAATCTGAACATCATCAAC 840  
 QY 841 CCAATATATGATTTGAACACACAGAGAGGCTCAGTCAAGAGGCTCAGAGGCTCAGAGAGT 900  
 Db 841 CCAATATATGATTTGAACACACAGAGAGGCTCAGTCAAGAGGCTCAGAGGCTCAGAGAGT 900  
 QY 901 ATCAGGTAAGTTGTTCAAACTGTGATGAGGCAATGTGACAAATACATCAATGCA 960  
 Db 901 ATCAGGTAAGTTGTTCAAACTGTGATGAGGCAATGTGACAAATACATCAATGCA 960

Db 901 ATCAGGTAGTTCGTTCCTTCAAACTTGATGTCGATGTCGACCCATGTCGCCAACAATCTCATGCGCA 960  
OY 961 GCTCATTCAGCATGAGAACAGCGATTATATCTACTACATAAAGACGAAATGATGATAGAAA 1020  
Db 961 GCTCATTCAGCATGAGAACAGCGATTATATCTACTACATAAAGACGAAATGATGATAGAAA 1020  
OY 1021 AGGCTGAATTCGTATATATAAAGCAACAGCCTGGCTTACGCAAGGAGCCACATACACAT 1080  
Db 1021 AGGCTGAATTCGTATATATAAAGCAACAGCCTGGCTTACGCAAGGAGCCACATACACAT 1080  
OY 1081 GGGCTGGAAGTATAGAAAACATGTATATAGCGGACTCCACGACACAGAAAAAGTAG 1140  
Db 1081 GGGCTGGAAGTATAGAAAACATGTATATAGCGGACTCCACGACACAGAAAAAGTAG 1140  
OY 1141 ATCTGAATGCTGATCCCTGTGTGAGAGAAAAGATAGCAATAGCAAGAACTGCCATGCT 1200  
Db 1141 ATCTGAATGCTGATCCCTGTGTGAGAGAAAAGATAGCAATAGCAAGAACTGCCATGCT 1200  
OY 1201 CAGAGAAATCCTAGAGATCTAGAGATGTTCTTGATTAACACTAAATAGCAGCATTCAGA 1260  
Db 1201 CAGAGAAATCCTAGAGATCTAGAGATGTTCTTGATTAACACTAAATAGCAGCATTCAGA 1260  
OY 1261 AAGTTAATGAGTGGTTTCCAGAGAGTGAATGTAAGTTAGTTCTGATGATCAGATGATG 1320  
Db 1261 AAGTTAATGAGTGGTTTCCAGAGAGTGAATGTAAGTTAGTTCTGATGATCAGATGATG 1320  
OY 1321 GGGAGTCTGAATCCAAATGCCAAAGTAGCTGATGATGTTGAGAGTTCTAATAGAGTAGTG 1380  
Db 1321 GGGAGTCTGAATCCAAATGCCAAAGTAGCTGATGATGTTGAGAGTTCTAATAGAGTAGTG 1380  
OY 1381 AATATTCGTGTTCTTCAGAGAAAATAGACTTACTGCGCAGTATCCATGATGAGGCTTTAA 1440  
Db 1381 AATATTCGTGTTCTTCAGAGAAAATAGACTTACTGCGCAGTATCCATGATGAGGCTTTAA 1440  
OY 1441 TATGTAAAAAGTAAAAAGATTCACTCAATCAGTGAAGATTAATTAAGACAAAATAT 1500  
Db 1441 TATGTAAAAAGTAAAAAGATTCACTCAATCAGTGAAGATTAATTAAGACAAAATAT 1500  
OY 1501 TTGGGAAAACCTATTCGGAAGAAGCGACCTGCCCAACTTAAGCATGATGAAGTAATC 1560  
Db 1501 TTGGGAAAACCTATTCGGAAGAAGCGACCTGCCCAACTTAAGCATGATGAAGTAATC 1560  
OY 1561 TAAATATAGGAGCAATTTGTACTGAGCCACAGATTAATCAAGAGGCTCCCTCACAAATA 1620  
Db 1561 TAAATATAGGAGCAATTTGTACTGAGCCACAGATTAATCAAGAGGCTCCCTCACAAATA 1620  
OY 1621 AATTAACCGTAAAAAGAGACCTTACATAGGCTTCACTCTGAGGATTTTTCAGAAAG 1680  
Db 1621 AATTAACCGTAAAAAGAGACCTTACATAGGCTTCACTCTGAGGATTTTTCAGAAAG 1680  
OY 1681 CAGATTTGGCAGTTCAAAAGACCTCTGAAATGATTAATCAGGGAACCTAACCAAGCGAGC 1740  
Db 1681 CAGATTTGGCAGTTCAAAAGACCTCTGAAATGATTAATCAGGGAACCTAACCAAGCGAGC 1740  
OY 1741 AGAATGGTCAAGTATGATATTTACTATAGTGTGATGAGATTAATAAACAAGAGTGAT 1800  
Db 1741 AGAATGGTCAAGTATGATATTTACTATAGTGTGATGAGATTAATAAACAAGAGTGAT 1800  
OY 1801 CTATTCAGATGAGAAAAATCCTTAACCAATAGATCACTCTGAAAAAGAAATCTGCTTCA 1860  
Db 1801 CTATTCAGATGAGAAAAATCCTTAACCAATAGATCACTCTGAAAAAGAAATCTGCTTCA 1860  
OY 1861 AAAGGAAAGCTGAACCTTAAGCAGCAGTATAGCAATATGGAAGTGAATTAATATACC 1920  
Db 1861 AAAGGAAAGCTGAACCTTAAGCAGCAGTATAGCAATATGGAAGTGAATTAATATACC 1920  
OY 1921 ACAATTCAAAAGCACTTAAGAAATAGCTGAGAGAGAGCTTCTTACCGACGATATTC 1980  
Db 1921 ACAATTCAAAAGCACTTAAGAAATAGCTGAGAGAGAGCTTCTTACCGACGATATTC 1980  
OY 1981 ATGGCTTGAGCTAGTATGATGAGAAATCTAAGCCCACTTAATTTGTAATTTGCAAA 2040  
Db 1981 ATGGCTTGAGCTAGTATGATGAGAAATCTAAGCCCACTTAATTTGTAATTTGCAAA 2040

OY 2041 TTGATAGTGTCTTCAGCAGTGAAGATTAAGAAAAAAGTACAAACCAATGCGCATCA 2100  
Db 2041 TTGATAGTGTCTTCAGCAGTGAAGATTAAGAAAAAAGTACAAACCAATGCGCATCA 2100  
OY 2101 GGCACAGCAGAAACCTTACAACTCATGGAAGGTAAAGAACTGCAACTGGAGCCAGAAAGA 2160  
Db 2101 GGCACAGCAGAAACCTTACAACTCATGGAAGGTAAAGAACTGCAACTGGAGCCAGAAAGA 2160  
OY 2161 GTACAAACCCAAATGAAACAGACAGTAAAGACATGACAGTATCTTTCCAGAGCTGA 2220  
Db 2161 GTACAAACCCAAATGAAACAGACAGTAAAGACATGACAGTATCTTTCCAGAGCTGA 2220  
OY 2221 AGTTAAACAAATGACACTGTTCTTTTACTAAGTGTCAAAATACCGATGACCTTAAGAT 2280  
Db 2221 AGTTAAACAAATGACACTGTTCTTTTACTAAGTGTCAAAATACCGATGACCTTAAGAT 2280  
OY 2281 TTGTCAATCTAGCCCTTCAGAGAGAGAGAAAAAGAGAAACTAGAAACAGTTAAAGTGT 2340  
Db 2281 TTGTCAATCTAGCCCTTCAGAGAGAGAGAAAAAGAGAAACTAGAAACAGTTAAAGTGT 2340  
OY 2341 CTAATTAATGCTGAAGACCCCAAGATCTCATGTTAAGTGAAGAAAGGCTTTTGCAACTG 2400  
Db 2341 CTAATTAATGCTGAAGACCCCAAGATCTCATGTTAAGTGAAGAAAGGCTTTTGCAACTG 2400  
OY 2401 AAAGATCTGTAGAGTAGCAGTATTTTCATTGCTACCTGATGATGATGACACTCAGG 2460  
Db 2401 AAAGATCTGTAGAGTAGCAGTATTTTCATTGCTACCTGATGATGATGACACTCAGG 2460  
OY 2461 AAAGATCTGTAGAGTAGCAGTATTTTCATTGCTACCTGATGATGATGACACTCAGG 2520  
Db 2461 AAAGATCTGTAGAGTAGCAGTATTTTCATTGCTACCTGATGATGATGACACTCAGG 2520  
OY 2521 GTGTGAGTCAAGTGTGACAGTATTTGAAAACCCCAAGGACTAATTCATGTTGTTCCAAAG 2580  
Db 2521 GTGTGAGTCAAGTGTGACAGTATTTGAAAACCCCAAGGACTAATTCATGTTGTTCCAAAG 2580  
OY 2581 ATTAATGAATGACACAGAAAGCTTAAGTATTCCTGAGCATGAAATTAACCAAGATC 2640  
Db 2581 ATTAATGAATGACACAGAAAGCTTAAGTATTCCTGAGCATGAAATTAACCAAGATC 2640  
OY 2641 GGGAAACAGCATAGAAATGGAAGAAAGTGAATGATGCTCAGTATTTGCAAGATACAT 2700  
Db 2641 GGGAAACAGCATAGAAATGGAAGAAAGTGAATGATGCTCAGTATTTGCAAGATACAT 2700  
OY 2701 TCAAGGTTTCAAAAGCGCAGTATTTGCTGTTTTCAAATCCAGAAATGAGAAAGAG 2760  
Db 2701 TCAAGGTTTCAAAAGCGCAGTATTTGCTGTTTTCAAATCCAGAAATGAGAAAGAG 2760  
OY 2761 AATGTGCAACATTTCTCTGCCACCTGCTGCTTAAAGAAACAAAGTCCAAAGCTCAT 2820  
Db 2761 AATGTGCAACATTTCTCTGCCACCTGCTGCTTAAAGAAACAAAGTCCAAAGCTCAT 2820  
OY 2821 TTGAATGTGAACAAAGAGAAATCAAGGAAAGATGAGTCTAATATCAAGCTGTAC 2880  
Db 2821 TTGAATGTGAACAAAGAGAAATCAAGGAAAGATGAGTCTAATATCAAGCTGTAC 2880  
OY 2881 AGACAGTTAATATATCATCTGAGGCTTCTGTGTTGTGACAGAAAGATAGCCAGTTGATA 2940  
Db 2881 AGACAGTTAATATATCATCTGAGGCTTCTGTGTTGTGACAGAAAGATAGCCAGTTGATA 2940  
OY 2941 ATGGCAAAATGATATCAAAAGAGGCTAGGTTTGTCTATCAATCTAGTTCAGAGGCA 3000  
Db 2941 ATGGCAAAATGATATCAAAAGAGGCTAGGTTTGTCTATCAATCTAGTTCAGAGGCA 3000  
OY 3001 ACAGAACTGAGCTATTAATCAATTAACATGACCTTTTCAAAAACCCATATGTAATAC 3060  
Db 3001 ACAGAACTGAGCTATTAATCAATTAACATGACCTTTTCAAAAACCCATATGTAATAC 3060  
OY 3061 CACCACTTTTCCATCAAGTCAATTTGTTAAACTAATGTAAGAAAAATCTGCTAGAGG 3120  
Db 3061 CACCACTTTTCCATCAAGTCAATTTGTTAAACTAATGTAAGAAAAATCTGCTAGAGG 3120

OY 3121 AAAACCTTGGAGAACATTCACATGTCACCTGAAAGAGAAATGGGAAATGAGAACATTCACAA 3180  
DB 3121 AAAACCTTGGAGAACATTCACATGTCACCTGAAAGAGAAATGGGAAATGAGAACATTCACAA 3180  
OY 3181 GTACAGGAGACAAATTAAGCCGTAAATACATTTAGAGAAAATGTTTTTAAAGAGCCACT 3240  
DB 3181 GTACAGGAGACAAATTAAGCCGTAAATACATTTAGAGAAAATGTTTTTAAAGAGCCACT 3240  
OY 3241 CAACCAATTAATTAAGAGTAGTTCACATTAATGAAGTGGGCTCCAGTATTAATGAAA 3300  
DB 3241 CAACCAATTAATTAAGAGTAGTTCACATTAATGAAGTGGGCTCCAGTATTAATGAAA 3300  
OY 3301 TAGGTTCCAGTGATGAAAAACATTCAGAGAACTAGTAGAAACAGAGGGCCAAAATTTGA 3360  
DB 3301 TAGGTTCCAGTGATGAAAAACATTCAGAGAACTAGTAGAAACAGAGGGCCAAAATTTGA 3360  
OY 3361 ATGCTATGCTTAGATTAGGGGTTTTGCACTGAGTCTATTAACAAAGTCTTCTGAAA 3420  
DB 3361 ATGCTATGCTTAGATTAGGGGTTTTGCACTGAGTCTATTAACAAAGTCTTCTGAAA 3420  
OY 3421 GTAATTTGAAGCATCTGAAATTAAGAAAGCAAGAAATATGAGAGTAGTTCAGACTTTA 3480  
DB 3421 GTAATTTGAAGCATCTGAAATTAAGAAAGCAAGAAATATGAGAGTAGTTCAGACTTTA 3480  
OY 3481 ATACAGATTTCTCTCCATATCTGATTTGATTAACATACTAGAACAGCCTATGSGAAGTAGTC 3540  
DB 3481 ATACAGATTTCTCTCCATATCTGATTTGATTAACATACTAGAACAGCCTATGSGAAGTAGTC 3540  
OY 3541 ATGATATCTCAGGTTGTTCTGAGACACCTGATGACCTGTTAGATGATGSGTAAATTAAGG 3600  
DB 3541 ATGATATCTCAGGTTGTTCTGAGACACCTGATGACCTGTTAGATGATGSGTAAATTAAGG 3600  
OY 3601 AAGATACTAGTTTTGCTGAAAAATGACATTTAAGAAAGTTTGTCTGTTTTTAAAGCAAAAGCG 3660  
DB 3601 AAGATACTAGTTTTGCTGAAAAATGACATTTAAGAAAGTTTGTCTGTTTTTAAAGCAAAAGCG 3660  
OY 3661 TCCAGAAAGAGAGACTTAGCAGAGAGTCTACGCCCTTTCACCATATACATTTGGCTCAGG 3720  
DB 3661 TCCAGAAAGAGAGACTTAGCAGAGAGTCTACGCCCTTTCACCATATACATTTGGCTCAGG 3720  
OY 3721 GTTACCCAGAGAGGGGCCAAGAAATTAAGTCTCAGAGAGAGACTTTATCTGTAGAGAGTG 3780  
DB 3721 GTTACCCAGAGAGGGGCCAAGAAATTAAGTCTCAGAGAGAGACTTTATCTGTAGAGAGTG 3780  
OY 3781 AAGAGCTTCCCTGCTTCCAAACACTTGTATTGTTGTTAAGTAAACAAATATACCTTCTCAGT 3840  
DB 3781 AAGAGCTTCCCTGCTTCCAAACACTTGTATTGTTGTTAAGTAAACAAATATACCTTCTCAGT 3840  
OY 3841 CTACTAGGCAATAGCAGCGTGTCTACGAGTGTCTGTCTAGAACACAGAGAGAGAAATTTAT 3900  
DB 3841 CTACTAGGCAATAGCAGCGTGTCTACGAGTGTCTGTCTAGAACACAGAGAGAGAAATTTAT 3900  
OY 3901 TATCATTTGAGAAATAGCTTAATGACTGACAGTACAGGTAATTTTGGCAAAAGGCAATCTC 3960  
DB 3901 TATCATTTGAGAAATAGCTTAATGACTGACAGTACAGGTAATTTTGGCAAAAGGCAATCTC 3960  
OY 3961 AGGAAATCACTCTTAGTGGAGAAAACAAATGTTCTGTAGCTTTTCTTCTCACAGTGCA 4020  
DB 3961 AGGAAATCACTCTTAGTGGAGAAAACAAATGTTCTGTAGCTTTTCTTCTCACAGTGCA 4020  
OY 4021 GTGAATTTGGAAGACTTACTGCAAAATACAAACCCAGAGATCTTCTCTGATTTGTTCTT 4080  
DB 4021 GTGAATTTGGAAGACTTACTGCAAAATACAAACCCAGAGATCTTCTCTGATTTGTTCTT 4080  
OY 4081 CCAAAACAATGAGGCAATCAGTCTGAAGCAAGGAGTTGGTCTGAGTGACAAAGAAATTTGG 4140  
DB 4081 CCAAAACAATGAGGCAATCAGTCTGAAGCAAGGAGTTGGTCTGAGTGACAAAGAAATTTGG 4140  
OY 4141 TTTTCAGATGATGAAGAAAGAGAAAGGCTTTGGAAGAAATTAATCAAGAGAGCAAAAGCA 4200  
DB 4141 TTTTCAGATGATGAAGAAAGAGAAAGGCTTTGGAAGAAATTAATCAAGAGAGCAAAAGCA 4200  
OY 4201 TGGATTTCAAACTTAGTGGAAGCAGCATCTGGGTGTGAGAGTGAAAACAAGCGCTCTGAGG 4260

DB 4201 TGGATTTCAAACTTAGTGGAAGCAGCATCTGGGTGTGAGAGTGAAAACAAGCGCTCTGAGG 4260  
OY 4261 ACTGCTCAGGGCTATTCCTCTCAGAGTGACATTTTAAACCATCTCAGCAGAGGATACCATG 4320  
DB 4261 ACTGCTCAGGGCTATTCCTCTCAGAGTGACATTTTAAACCATCTCAGCAGAGGATACCATG 4320  
OY 4321 AACATTAACCTGATTAAGCTCCAGCAGAAATGGCTCACTAGAAAGCTGTTTGAACACG 4380  
DB 4321 AACATTAACCTGATTAAGCTCCAGCAGAAATGGCTCACTAGAAAGCTGTTTGAACACG 4380  
OY 4381 ATGGAGCCAGCTTCTPAACAGCTACCTTCCATTAATAGTAGCTCTTCTGCTTGAAG 4440  
DB 4381 ATGGAGCCAGCTTCTPAACAGCTACCTTCCATTAATAGTAGCTCTTCTGCTTGAAG 4440  
OY 4441 ACCTGCGAAATCCGAGCAAAAGCAGATCAGAAAAAGAGATTAATTAATCTCAGCAAAAGTA 4500  
DB 4441 ACCTGCGAAATCCGAGCAAAAGCAGATCAGAAAAAGAGATTAATTAATCTCAGCAAAAGTA 4500  
OY 4501 GTGAATACCCATTAAGCCAGATCCAGAAAGGCTTCTGCTGCAAGATTTGAGTGTCTG 4560  
DB 4501 GTGAATACCCATTAAGCCAGATCCAGAAAGGCTTCTGCTGCAAGATTTGAGTGTCTG 4560  
OY 4561 CAGATTAAGTTCTTACAGTAAATTAAGAACCCAGAGTGAAGAGTCAATCCCTTTCTAAT 4620  
DB 4561 CAGATTAAGTTCTTACAGTAAATTAAGAACCCAGAGTGAAGAGTCAATCCCTTTCTAAT 4620  
OY 4621 GCCCATCATTAATGATGATGAGTGTGATGACATGACAGTGTCTGAGAGTCTTCAAAATTAAGA 4680  
DB 4621 GCCCATCATTAATGATGATGAGTGTGATGACATGACAGTGTCTGAGAGTCTTCAAAATTAAGA 4680  
OY 4681 ACTACCCATCTCAAGAGAGAGCTCTTAAGGTTGTGATGTGAGAGAGCAACAGCTGGAAG 4740  
DB 4681 ACTACCCATCTCAAGAGAGAGCTCTTAAGGTTGTGATGTGAGAGAGCAACAGCTGGAAG 4740  
OY 4741 AGTCTGGGCCACACAGATTTTGAAGCAAACTTACTTGTCCCAAGGCAAGATCTAGAGGAA 4800  
DB 4741 AGTCTGGGCCACACAGATTTTGAAGCAAACTTACTTGTCCCAAGGCAAGATCTAGAGGAA 4800  
OY 4801 CCCCTTACCTGGAATCTGGAATCAGGCTCTTCTGTAAGGACCTGGAATGATCTTCTGAG 4860  
DB 4801 CCCCTTACCTGGAATCTGGAATCAGGCTCTTCTGTAAGGACCTGGAATGATCTTCTGAG 4860  
OY 4861 AAGACAGAGGCCAGAGATCAGTCTGTGTGGCAACATACCATTTTAACTCTGCAATGA 4920  
DB 4861 AAGACAGAGGCCAGAGATCAGTCTGTGTGGCAACATACCATTTTAACTCTGCAATGA 4920  
OY 4921 AAGTTCCTCCCAATTAAGATTTGCAAGATCTGCCCCAGAGGCTCAGCTCTCATACTAGTG 4980  
DB 4921 AAGTTCCTCCCAATTAAGATTTGCAAGATCTGCCCCAGAGGCTCAGCTCTCATACTAGTG 4980  
OY 4981 ATACTGCTGGGTAAATGCAATGGAAGAAAGTGTGAGCAAGGAGAGGAGCAGAAATTAACG 5040  
DB 4981 ATACTGCTGGGTAAATGCAATGGAAGAAAGTGTGAGCAAGGAGAGGAGCAGAAATTAACG 5040  
OY 5041 CTTCAACAGAAAGGCTCAACAAAAGAAATGTCAATGTGTGTCTGAGCCAGCCAGAG 5100  
DB 5041 CTTCAACAGAAAGGCTCAACAAAAGAAATGTCAATGTGTGTCTGAGCCAGCCAGAG 5100  
OY 5101 AATTATGCTGCTGTACAAAGTTTGCAGAAACACACATCACTTAATCTAATTA 5160  
DB 5101 AATTATGCTGCTGTGTACAAAGTTTGCAGAAACACACATCACTTAATCTAATTA 5160  
OY 5161 CTGAAGAGACTACTCATGTTGTTATGAAAACAGATCTGAGTTTGTGTGAACGACAC 5220  
DB 5161 CTGAAGAGACTACTCATGTTGTTATGAAAACAGATCTGAGTTTGTGTGAACGACAC 5220  
OY 5221 TGAATATTTTCTAGCAATTTGCGGAGAGAAAGGAGTATTAATCTGAGTGTGAGCC 5280  
DB 5221 TGAATATTTTCTAGCAATTTGCGGAGAGAAAGGAGTATTAATCTGAGTGTGAGCC 5280  
OY 5281 AGTCTATTTAAAGAAAGAAATGCTGAATGAGCATGATTTTGAAGTCAAGAGAGATGTGG 5340

Db 5281 ACTCTTTAAGAAAGAAAAATGCTGATGAGCATGATTTTGAATGACAGAGATGTGG 5340  
QY 5341 TCATGGAAGAAACCCACCAAGGTCCAAAGCGAGAGAGATCCAGAGAGAAATCT 5400  
Db 5341 TCATGGAAGAAACCCACCAAGGTCCAAAGCGAGAGAGATCCAGAGAGAAATCT 5400  
QY 5401 TCAGGGGGCTAGAAATCTGTGCTATGAGGCCCTTCACCAACATGCCACAGATCAACTGG 5460  
Db 5401 TCAGGGGGCTAGAAATCTGTGCTATGAGGCCCTTCACCAACATGCCACAGATCAACTGG 5460  
QY 5461 AATGATGCTACAGCTGTGTGGTCTCTGTGTGTAAGAGAGCTTTATCATTCACCTTG 5520  
Db 5461 AATGATGCTACAGCTGTGTGGTCTCTGTGTGTAAGAGAGCTTTATCATTCACCTTG 5520  
QY 5521 GCACAGGTGTCACCAATTTGTTGTGTCAGCCACATGCTGGACAGAGAGATGGCT 5580  
Db 5521 GCACAGGTGTCACCAATTTGTTGTGTCAGCCACATGCTGGACAGAGAGATGGCT 5580  
QY 5581 TCCATGCAATTTGGGAGATGTGTGAGGCACTGTGTGACCCGAGAGTGGTGTGGACA 5640  
Db 5581 TCCATGCAATTTGGGAGATGTGTGAGGCACTGTGTGACCCGAGAGTGGTGTGGACA 5640  
QY 5641 GTGTACACTCTACCAAGTGGCAGAGAGCTGGACACCTACCTGATACCCCGAGATCCCCACA 5700  
Db 5641 GTGTACACTCTACCAAGTGGCAGAGAGCTGGACACCTACCTGATACCCCGAGATCCCCACA 5700  
QY 5701 GCCACTACTGA 5711  
Db 5701 GCCACTACTGA 5711

## RESULT 8

US-09-911-904-127  
; Sequence 127, Application US/09911904  
; Publication No. US20030096234A1  
; GENERAL INFORMATION:  
; APPLICANT: Farr, Spencer B.  
; APPLICANT: Pickett, Gavin G.  
; APPLICANT: Neft, Robin Eileen  
; APPLICANT: Dunn, II, Robert Thomas  
; TITLE OF INVENTION: CANINE TOXICITY GENES  
; FILE REFERENCE: 400742000200  
; CURRENT APPLICATION NUMBER: US/09/911,904  
; PRIOR FILING DATE: 2002-04-09  
; PRIOR APPLICATION NUMBER: US 60/220,057  
; PRIOR FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 386  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 127  
; LENGTH: 499  
; TYPE: DNA  
; ORGANISM: Canis familiaris  
US-09-911-904-127

Query Match 6.4%; Score 364.6; DB 9; Length 499;  
Best Local Similarity 86.7%; Pred. No. 1.9e-82;  
Matches 419; Conservative 0; Mismatches 49; Indels 15; Gaps 1;

QY 5229 TTTCTGGAATTCGGGAGGAAATGGGTAGTATTTCGCGGTGACCCAGTCTATT 5288  
Db 1 TTTCTGGGTATTCAGAGGAAATGGGTAGTATTTCGCGGTGACCCAGTCTATT 60  
QY 5289 AAGAAAGAAATGCTGATGAGCATGATTTTGAAGTACAGAGATGTGGTCAATGGA 5348  
Db 61 AAGAAAGAAATGCTGATGAGCATGATTTTGAAGTACAGAGATGTGGTCAATGGA 120  
QY 5349 AGAAGACCAAGTCCAAAGCGAGAGAGATCCAG-----GACAGA 5393  
Db 121 AGAAGTACCAAGTCCAAAGCGAGAGAGATCCAGAGACAGAGATCCCAAGACAGA 180  
QY 5394 AAGATCTTACAGGGGCTAGAAATCTGTGCTATGAGGCCCTTCACCAACATGCCACAGAT 5453  
Db 181 AAGATCTTACAGGGGCTAGAAATCTGTGCTATGAGGCCCTTCACCAACATGCCACAGAT 240

QY 5454 CAACGTGAATGATGATACAGCTGTGTGCTGTCTGTGTGAGAGACCTTCATCATTC 5513  
Db 241 CAATTTAGATGATGATGATACAGCTGTGTGCTGTCTGTGTGAGAGACCTTCATTC 300  
QY 5514 ACCCTTGGCAGAGTGTCCACCAATTTGTTGTGTCAGCCAGATGCTGGACAGAGAC 5573  
Db 301 ACCCTTGGCAGAGTGTCCACCAATTTGTTGTGTCAGCCAGATGCTGGACAGAGAC 360  
QY 5574 AATGCTTCATCATTCATTTGGGAGATGTGTGAGGACCTGTGTGACCCGAGAGTGGCTG 5633  
Db 361 AATGCTTCATCATTCATTTGGGAGATGTGTGAGGACCTGTGTGACCCGAGAGTGGCTG 420  
QY 5634 TTGACAGTGTACACTCTACCAAGTGGCAGAGAGCTGGACACCTGATACCCAGATC 5693  
Db 421 CTGACAGTGTACACTCTACCAAGTGGCAGAGAGCTGGACACCTGATACCCAGATC 480  
QY 5694 CCC 5696  
Db 481 CCC 483

## RESULT 9

US-09-864-761-4552  
; Sequence 4552, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aecmca-x-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263,6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117



```

1 SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
2 SEQ ID NO 4552
3 LENGTH: 424
4 TYPE: DNA
5 ORGANISM: Homo sapiens
6
7 FEATURE:
8 OTHER INFORMATION: MAP TO L78933.1
9 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL
10 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL
11 OTHER INFORMATION: EXPRESSED IN LONG, SIGNAL = 2
12 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL
13 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.3
14 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL =
15 OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.9
16 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9
17 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.4
18 OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.3
19
20 US-09-864 -761 -4552

```

Query Match	3.1%	Score 175.4	DB 10	length 424
Best Local Similarity	94.3%	Pred. No. 2.7e-34		
Matches 182	Conservative	0	Mismatches 11	Indels 0
				Gaps 0

QY	4303	GGCAGAGAGGATACCATGCACATATTAACCTGATAAAGCTCCACAGAGAAATGGCTGAACAG	4362
Db	232	AGCAGAGAGGATACCATGCACATATTAACCTGATAAAGCTCCACAGAGAAATGGCTGAACAG	291
QY	4363	AAGCTGTGTATGAACAGCATGGGAGCCAGCCTTCAACAGCTACCTTCATCATTAATG	4422
Db	292	AAGCTGTGTATGAACAGCATGGGAGCCAGCCTTCAACAGCTACCTTCATCATTAATG	351
QY	4423	ACTCTTTCGCCCTTGAGGACCTGGGAAATCCAGAACAAAGCATCAGAAAAAGCAGTAT	4482
Db	352	ACTCTTTCGCCCTTGAGGACCTGGGAAATCCAGAACAAAGCATCAGAAAAAGTGTGT	411
QY	4483	TAACTTCACAGAA	4495
Db	412	ATTGTTGGCCAAA	424

RESULT 10  
US-09-864-761-21299  
Sequence 21299, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aecm1ca-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665

```
:
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00666
:
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00666
:
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00666
:
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00666
:
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
:
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
:
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 9/608,408
:
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
:
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ. NO. 4011
```

```

; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1.1
; SEQ ID NO 21299
; LENGTH: 147

```

```

? ORGANISM: Homo sapiens
? FEATURE:
? OTHER INFORMATION: MAP TO L78833.1
? OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.4
? OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
? OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2
? OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.1
? OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.3
? OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2
? OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.9
? OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9
? OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.4
? OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.3
? OTHER INFORMATION: EST HUMAN HIT: A0125312.1, EVALUE 2.00e-77
? OTHER INFORMATION: SWISSPROT HIT: P38398, EVALUE 4.00e-12
? OTHER INFORMATION: NT HIT: g11424969, EVALUE 1.00e-77
? OS:99-864-761-21299

```

	Query Match	2.6%	Score 147	DB 10	Length 147
	Best Local Similarity	100.0%	Pred. NO. 2.3e-27		
	Matches 147	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	4330	TGATAAGCTCCAGCAGCAAAATGCGTGAACATGAGAGCTGTGTTAAGAACAGCATGGAGGCC	4389		
Db	1	TGATAAGCTCCAGCAGCAAAATGCGTGAACATGAGAGCTGTGTTAAGAACAGCATGGAGGCC	60		
QY	4390	AGCCTTCTAACAGCTACCCCTTCATCATTAAGTGAAGCTTTGCGCCTTGAAGAGCTTCGAA	4445		
Db	61	AGCCTTCTAACAGCTACCCCTTCATCATTAAGTGAAGCTTTGCGCCTTGAAGAGACTTCGAA	120		
QY	4450	ATCCAGAACAAAGCAGCATCAGAAAAAG	4476		
Db	121	ATCCAGAACAAAGCAGCATCAGAAAAAG	147		

```

RESULT 11
US-09-818--875-653
: Sequence 653, Application US/09818875
: Publication No. US20030051270A1
: GENERAL INFORMATION:
: APPLICANT: Kmiec, Eric B.
: APPLICANT: Gamper, Howard B.
: APPLICANT: Rice, Michael C.
: TITLE OF INVENTION: Targeted Chromosomal Genomic Alterations with Modified Single
: TITLE OF INVENTION: Stranded Oligonucleotides
: FILE REFERENCE: Napro-4
: CURRENT APPLICATION NUMBER: US/09/818,875
: CURRENT FILING DATE: 2001-03-27
: PRIOR APPLICATION NUMBER: US 60/192,176
: PRIOR FILING DATE: 2000-03-27
: PRIOR APPLICATION NUMBER: US 60/192,179

```



```

? PRIOR FILING DATE: 2000-03-27
? PRIOR APPLICATION NUMBER: US 60/208,538
? PRIOR FILING DATE: 2000-06-01
? PRIOR APPLICATION NUMBER: US 60/244,989
? PRIOR FILING DATE: 2000-10-30
? NUMBER OF SEQ ID NOS: 4385
? SOFTWARE: Friedman macro Naprod
? SEQ ID NO 653
? LENGTH: 121
? TYPE: DNA
? ORGANISM: Homo sapiens
US-09-818-875-653

```

Query Match	2.1%	Score 121	-DB 9	Length 121
Best Local Similarity	100.0%	Pred. No.	8	6e-21
Matches 121	Conservative 0	Mismatches	0	Gaps 0

QY	62	CTGCGGCTCAGAGAGGCGCTTCACCCCTGCTCCTGGTAAAGTTCATTTGGAACAGAAAGAAAT	121
		1	CTGCGGCTCAGAGAGGCGCTTCACCCCTGCTCCTGGTAAAGTTCATTTGGAACAGAAAGAAAT
Db			60
QY	122	GGATTATATCGCTTCGCGCTGGAAGAAAGTACCAAAATGTCATTTAATGCTATGCAGAAAT	181
			120
Db	61	GGATTATATCGCTTCGCGCTGGAAGAAAGTACCAAAATGTCATTTAATGCTATGCAGAAAT	120
QY			
	182	C	182
QY			
	121	C	121
Db			

```

RESULT 12
US-09-818-875-654/C
; Sequence 654, Application US/09818875
; Publication No. US20030051270A1
; GENERAL INFORMATION:
; APPLICANT: Kmiec, Eric B.
; APPLICANT: Gamper, Howard B.
; APPLICANT: Rice, Michael C.
; TITLE OF INVENTION: Targeted Chromosomal Genomic Alterations with Modified Single
; FILE REFERENCE: Napro-4
; CURRENT APPLICATION NUMBER: US/09/818,875
; CURRENT FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: US 60/192,176
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: US 60/192,179
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: US 60/208,538
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/244,989
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 4385
; SOFTWARE: Fliedman macro Napro4
; SEQ ID NO 654
; LENGTH: 121
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-818-875-654

```

	Query Match	2.1%	Score 121	DB 9	Length 121	
	Best Local Similarity	100.0%	Pred. NO. 8.6e-21			
	Matches 121	Conservative 0	Mismatches 0	Indels 0	Gaps 0	
QY	62	CTGGCGCTCAGAGAGCCCTTCACCCCTCTGCTCGGGTAAGTCATTCGGAACAAGAAAT	121			
Db	121	CTGGCGCTCAGAGAGCCCTTCACCCCTCTGCTCGGGTAAGTCATTCGGAACAAGAAAT	62			
QY	122	GGATTATCTGCTCTTGGCGGTTGAAGAAGTCAAAATCTCATTTAATGCTATGACAGAAAT	181			
Db	61	GGATTATCTGCTCTTGGCGGTTGAAGAAGTCAAAATCTCATTTAATGCTATGACAGAAAT	2			
QY	182	C	182			

Db 1 C 1

RESULT 13  
US-09-818--875--657  
; Sequence 657, Application US/09818875  
; Publication No. US20030051270A1  
; GENERAL INFORMATION:

Query Match	2.1%	Score 121	DB 91	length 121
Best Local Similarity	100.0%	Pred. No.	8.6e-21	
Matches 121	Conservative 0	Mismatches 0	Indels 0	Gaps 0

QY 91 CTGGGTAAAGTCATCTGGAAACAGAAAGAAATGCATTTATCGCTTCGCTTGAACAG 150  
Db 1 CTGGGTAAAGTCATCTGGAAACAGAAAGAAATGCATTTATCGCTTCGCTTGAACAG 60  
QY 151 TACAAATGTCATTAATGCTATGCGAAGAAATCTTAGAGTCCCATCGTCTGGAGTTGA 210  
Db 61 TACAAATGTCATTAATGCTATGCGAAGAAATCTTAGAGTCCCATCTGTCTGGAGTTGA 120  
QY 211 T 211  
Db 121 T 121

RESULT 14  
US-09-818-875-658/C  
Sequence: 658, Application US/09818875  
Publication No. US20030051270A1  
GENERAL INFORMATION:  
APPLICANT: kmiec, Eric B.  
APPLICANT: Gamper, Howard B.  
APPLICANT: Rice, Michael C.  
TITLE OF INVENTION: Targeted Chromosomal Genomic Alterations with Modified Single  
TITLE OF INVENTION: Stranded Oligonucleotides  
FILE REFERENCE: Napro-4  
CURRENT APPLICATION NUMBER: US/09/818,875  
CURRENT FILING DATE: 2001-03-27  
PRIOR APPLICATION NUMBER: US 60/192,176  
PRIOR FILING DATE: 2000-03-27  
PRIOR APPLICATION NUMBER: US 60/192,179  
PRIOR FILING DATE: 2000-03-27  
PRIOR APPLICATION NUMBER: US 60/208,538  
PRIOR FILING DATE: 2000-06-01  
PRIOR APPLICATION NUMBER: US 60/244,989  
PRIOR FILING DATE: 2000-10-30  
NUMBER OF SEQ ID NOS: 4385  
SOFTWARE: Friedman macro Napro4  
SEQ ID NO 658

; LENGTH: 121  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-818-875-658

Query Match 2.1%; Score 121; DB 9; Length 121;  
Best Local Similarity 100.0%; Pred. No. 8.6e-21;  
Matches 121: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 91 CTGGTAAGTTCATTGGAAGAAAGAAATGATTTATCTGCTCTTCCGTTGAAGAG 150  
|  
Db 121 CTGGTAAGTTCATTGGAAGAAAGAAATGATTTATCTGCTCTTCCGTTGAAGAG 62  
  
QY 151 TACAAATGTCATTATGCTATGCAAAAATCTTAGAGTGCCATCTGCTGAGTTGA 210  
|  
Db 61 TACAAATGTCATTATGCTATGCAAAAATCTTAGAGTGCCATCTGCTGAGTTGA 2  
  
QY 211 T 211  
|  
Db 1 T 1

## RESULT 15

US-09-818-875-661  
; Sequence 661, Application US/09818875  
; Publication No. US20030051270A1  
; GENERAL INFORMATION:  
; APPLICANT: Kmiec, Eric B.  
; APPLICANT: Gamper, Howard B.  
; APPLICANT: Rice, Michael C.  
; TITLE OF INVENTION: Targeted Chromosomal Genomic Alterations with Modified Single  
; FILE REFERENCE: Napro-4  
; CURRENT APPLICATION NUMBER: US/09/818,875  
; PRIOR FILING DATE: 2001-03-27  
; PRIOR APPLICATION NUMBER: US 60/192,176  
; PRIOR FILING DATE: 2000-03-27  
; PRIOR APPLICATION NUMBER: US 60/192,179  
; PRIOR FILING DATE: 2000-03-27  
; PRIOR APPLICATION NUMBER: US 60/208,538  
; PRIOR FILING DATE: 2000-06-01  
; PRIOR APPLICATION NUMBER: US 60/244,989  
; PRIOR FILING DATE: 2000-10-30  
; NUMBER OF SEQ ID NOS: 4385  
; SOFTWARE: Friedman macro Napro4  
; SEQ ID NO 661  
; LENGTH: 121  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-818-875-661

Query Match 2.1%; Score 121; DB 9; Length 121;  
Best Local Similarity 100.0%; Pred. No. 8.6e-21;  
Matches 121: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 ATGATTTATCTGCTCTTGGCTTGAGAGTACAAAATGCTAATGCTATGCAGAAA 179  
|  
Db 1 ATGATTTATCTGCTCTTGGCTTGAGAGTACAAAATGCTAATGCTATGCAGAAA 60  
  
QY 180 ATCTTAGAGTCTCCATCTGCTGAGATTGATCAAGAACTGCTCCACAAAGTGTGAC 239  
|  
Db 61 ATCTTAGAGTCTCCATCTGCTGAGATTGATCAAGAACTGCTCCACAAAGTGTGAC 120  
  
QY 240 C 240  
|  
Db 121 C 121

Search completed: June 27, 2003, 05:24:49  
Job time : 763.849 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 26, 2003, 12:33:08 ; Search time 1078.23 Seconds  
(without alignments)  
11673.202 Million cell updates/sec

Title: US-09-734-672-3\_COPY\_120\_5708

Perfect score: 5589  
Sequence: 1 ATGCATTATTCGCTCTCG.....AGATCCCGACAGCCACTAC 5589

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :  
1: N\_Geneseq\_101002.\*  
2: /SID2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:\*  
3: /SID2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:\*  
4: /SID2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:\*  
5: /SID2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:\*  
6: /SID2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:\*  
7: /SID2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:\*  
8: /SID2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:\*  
9: /SID2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:\*  
10: /SID2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:\*  
11: /SID2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:\*  
12: /SID2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:\*  
13: /SID2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:\*  
14: /SID2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:\*  
15: /SID2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:\*  
16: /SID2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:\*  
17: /SID2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:\*  
18: /SID2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:\*  
19: /SID2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:\*  
20: /SID2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:\*  
21: /SID2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:\*  
22: /SID2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:\*  
23: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:\*  
24: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:\*  
25: /SID2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5587.4	100.0	5711	19	AAV46449
2	5587.4	100.0	5711	21	AAC60795
3	5587.4	100.0	5711	17	AAAT17493
4	5587.4	100.0	5711	17	AAAT17495
5	5587	100.0	5711	19	AAV46463
6	5587	100.0	5711	19	AAV46464
7	5587	100.0	5711	19	AAV46458
8	5587	100.0	5711	19	AAV46459
9	5587	100.0	5711	19	AAV46460

10	5587	100.0	5711	19	AAV46461	Human BRCA1 omi2 p
11	5587	100.0	5711	19	AAV46462	Human BRCA1 omi2 p
12	5585.8	99.9	5712	18	AAAT84840	Human breast and o
13	5585.8	99.9	5712	21	AAAT87995	BRCA1 gene sequenc
14	5585.8	99.9	5914	17	AAAT32601	BRCA1, breast and
15	5585.8	99.9	5914	17	AAAT17438	BRCA1 coding seque
16	5585.8	99.9	5914	17	AAAT18310	BRCA1, breast and
17	5585.8	99.9	14533	23	AAAT58883	DNA encoding novel
18	5584.2	99.9	5711	17	AAAT42024	BRCA1 allele #8403
19	5584.2	99.9	5711	17	AAAT42030	BRCA1 allele #7408
20	5584.2	99.9	5711	18	AAAT70067	Mutant BRCA1 allel
21	5584.2	99.9	5711	18	AAAT70073	Mutant BRCA1 allel
22	5584.2	99.9	5711	19	AAV60569	Tumorigenic BRCA1
23	5584.2	99.9	5711	19	AAV60569	Tumorigenic BRCA1
24	5584.2	99.9	5914	17	AAAT17445	Mutated BRCA1 codi
25	5584.2	99.9	5914	17	AAAT17446	Mutated BRCA1 codi
26	5584.2	99.9	5914	17	AAAT17449	Mutated BRCA1 codi
27	5584.2	99.9	5914	17	AAAT17453	Mutated BRCA1 codi
28	5584.2	99.9	5914	17	AAAT17457	Mutated BRCA1 codi
29	5584.2	99.9	5914	17	AAAT17462	Mutated BRCA1 codi
30	5584.2	99.9	5914	17	AAAT17466	Mutated BRCA1 codi
31	5584.2	99.9	5914	17	AAAT17468	Mutated BRCA1 codi
32	5584.2	99.9	5914	17	AAAT17471	Mutated BRCA1 codi
33	5584.2	99.9	5914	17	AAAT17476	Mutated BRCA1 codi
34	5584.2	99.9	5914	17	AAAT17478	Mutated BRCA1 codi
35	5584.2	99.9	5914	17	AAAT17479	Mutated BRCA1 codi
36	5584.2	99.9	5914	17	AAAT17481	Mutated BRCA1 codi
37	5584.2	99.9	5914	17	AAAT17483	Mutated BRCA1 codi
38	5584.2	99.9	5914	17	AAAT17484	Mutated BRCA1 codi
39	5584.2	99.9	5914	17	AAAT17486	Mutated BRCA1 codi
40	5584.2	99.9	5914	17	AAAT17489	Mutated BRCA1 codi
41	5584.2	99.9	5914	17	AAAT17490	Mutated BRCA1 codi
42	5584.2	99.9	5914	17	AAAT17491	Mutated BRCA1 codi
43	5584.2	99.9	5914	17	AAAT17492	Mutated BRCA1 codi
44	5584.2	99.9	5914	17	AAAT17494	Mutated BRCA1 codi
45	5584.2	99.9	5914	17	AAAT17496	Mutated BRCA1 codi

## ALIGNMENTS

RESULT 1	AAV46449	standard; cDNA; 5711 BP.
ID	AAV46449	
XX	AAV46449;	
AC	18-NOV-1998	(first entry)
XX		
DT	Human BRCA1 omi2 cDNA.	
XX		
DE	BRCA1; omi2; human; breast and ovarian cancer predisposing gene;	
KW	polymorphism; susceptibility; anti-oncogene; tumour suppressor;	
KW	chromosome 17q; ss.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	120..5711
FT		/*tag= a
XX		/product= "BRCA1 omi2 protein"
PN	US5750400-A.	
XX		
PD	12-MAY-1998.	
XX		
PF	12-FEB-1997;	97US-0798691.
XX		
PR	12-FEB-1996;	96US-0598591.
XX	12-FEB-1997;	97US-0798691.
XX		
PA	(ONCO-) ONCORMED INC.	
XX		



1741 AAAAGAAAGCTGAACCTATTAAGCAGCAGTATTAAGCAATATGGAAGTCAATTAATATATC 1800  
1860 AAAAGAAAGCTGAACCTATTAAGCAGCAGTATTAAGCAATATGGAAGTCAATTAATATATC 1919  
1801 CACAAATTCAAAAGCAGCTTAAAAAGATAGGCTGAGAGAGAGAGCTTCTACAGGCAATAT 1860  
1920 CACAAATTCAAAAGCAGCTTAAAAAGATAGGCTGAGAGAGAGAGCTTCTACAGGCAATAT 1979  
1861 CATGGCCTGAAGTACTAGTACAGTAAATCTAAGCCCACTAATGTGTCTGATTAATGCA 1920  
1980 CATGGCCTGAAGTACTAGTAAATCTAAGCCCACTAATGTGTCTGATTAATGCA 2039  
1921 ATTGATAGTGTCTAGCAGTAAAGATTAAGAAAAAAGTACAACTAATGCGCAGTC 1980  
2040 ATTGATAGTGTCTAGCAGTAAAGATTAAGAAAAAAGTACAACTAATGCGCAGTC 2099  
1981 AGGCACAGCAGAACTTACACTCATGTAAGGTAAAGAACTGCAACTGGAGCCAGGAAG 2040  
2100 AGGCACAGCAGAACTTACACTCATGTAAGGTAAAGAACTGCAACTGGAGCCAGGAAG 2159  
2041 AGTAACAGCCAAATGAGACAGCAAGTAAAGACATGAGATCTTCCAGAGCTG 2100  
2160 AGTAACAGCCAAATGAGACAGCAAGTAAAGACATGAGATCTTCCAGAGCTG 2219  
2101 AAGTTAACAATGACACCTGTTCTTTTACTAAGTGTTCAAATACAGTAACTTAAGAA 2160  
2220 AAGTTAACAATGACACCTGTTCTTTTACTAAGTGTTCAAATACAGTAACTTAAGAA 2279  
2161 TTTGTCAATCTAGCCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220  
2280 TTTGTCAATCTAGCCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2339  
2221 TCTAATATGCTGAAGACCCCAAGATCTCATGTGTAGTGAAGAGGTTTTCGAAGT 2280  
2340 TCTAATATGCTGAAGACCCCAAGATCTCATGTGTAGTGAAGAGGTTTTCGAAGT 2399  
2281 GAAAGATCTGTAGAGATGAGAGATATTCATGCTGCTGATGATTAATGAGCAGTACG 2340  
2400 GAAAGATCTGTAGAGATGAGAGATATTCATGCTGCTGATGATTAATGAGCAGTACG 2459  
2341 GAAAGATCTGTAGAGATGAGAGATATTCATGCTGCTGATGATTAATGAGCAGTACG 2400  
2460 GAAAGATCTGTAGAGATGAGAGATATTCATGCTGCTGATGATTAATGAGCAGTACG 2519  
2401 TGTGTAGTCACTGTGACAGATTTGAAACCCCAAGAGAGATTAATGATGTTGTTCCAA 2460  
2520 TGTGTAGTCACTGTGACAGATTTGAAACCCCAAGAGAGATTAATGATGTTGTTCCAA 2579  
2461 GATTAATGAAATGACACAGAGGCTTTAAGTATCCATTGGGACATGAAGTTAACCAAGT 2520  
2580 GATTAATGAAATGACACAGAGGCTTTAAGTATCCATTGGGACATGAAGTTAACCAAGT 2639  
2521 GGGGAAACAGAGATGAAATGGAAGAAATGAACTGATGCTCAGATTTGCAAGATTA 2580  
2640 GGGGAAACAGAGATGAAATGGAAGAAATGAACTGATGCTCAGATTTGCAAGATTA 2699  
2581 TTTCAAGGTTTCAAAAGCAGCAGTATTTGCTGTTTCAATTCAGAGAAATGAGAGAG 2640  
2700 TTTCAAGGTTTCAAAAGCAGCAGTATTTGCTGTTTCAATTCAGAGAAATGAGAGAG 2759  
2641 GAATGTGCAACATTTCTGCCCCACTGCGGCTCTTAAGAAACAAAGTCCAAAGTCACT 2700  
2760 GAATGTGCAACATTTCTGCCCCACTGCGGCTCTTAAGAAACAAAGTCCAAAGTCACT 2819  
2701 TTTGATATGGAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2760  
2820 TTTGATATGGAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2879  
2761 CAGACAGTAAATATCACTGAGGCTTTCTGTTGTTGTCAGAGAGAGATTAAGCAGTTGAT 2820  
2880 CAGACAGTAAATATCACTGAGGCTTTCTGTTGTTGTCAGAGAGAGATTAAGCAGTTGAT 2939  
2821 AATGCCAAATGTAGTATCAAGAGAGGCTCTAGGTTTGTCTATCATCTCAGTTCAAGAGC 2880

2940 AATGCCAAATGTAGTATCAAGAGAGGCTCTAGGTTTGTCTATCATCTCAGTTCAAGAGC 2999  
2881 AAGGAACCTGACTCATTTCTCCAAATTAACATGAGCTTTTACAAAACCATATGCTATA 2940  
3000 AAGGAACCTGACTCATTTCTCCAAATTAACATGAGCTTTTACAAAACCATATGCTATA 3059  
2941 CCACACTTTTCCATCAAGTATTTGTTAAACCTAATGTAAAGAAATGCTAGAG 3000  
3060 CCACACTTTTCCATCAAGTATTTGTTAAACCTAATGTAAAGAAATGCTAGAG 3119  
3001 GAAACTTTGAGAGACATTCATGTCACCTGAAAGAGAAATGGAATGAGAACATTTCA 3060  
3120 GAAACTTTGAGAGACATTCATGTCACCTGAAAGAGAAATGGAATGAGAACATTTCA 3179  
3061 AGTACAGTGAACCAATTAAGCCGTAATTAACATTAAGAGAAATGTTTAAAGAGCCAGC 3120  
3180 AGTACAGTGAACCAATTAAGCCGTAATTAACATTAAGAGAAATGTTTAAAGAGCCAGC 3239  
3121 TCAGCAATATTAATGAGTATGAGTTCAGTAAATGAAGTGGGCTCCAGTATTAATGA 3180  
3240 TCAGCAATATTAATGAGTATGAGTTCAGTAAATGAAGTGGGCTCCAGTATTAATGA 3299  
3181 ATAGGTTCCAGTATGAGAGAAACATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3240  
3300 ATAGGTTCCAGTATGAGAGAAACATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3359  
3241 AATGCTATGCTTGAATTAAGGAGGTTTTCGACACCTGAGTCTATTAACAAATCTTCTGCA 3300  
3360 AATGCTATGCTTGAATTAAGGAGGTTTTCGACACCTGAGTCTATTAACAAATCTTCTGCA 3419  
3301 AGTAATTTGTAAGCATCTCTGAATTAATAAAGCAAGTATTAAGTATGAGTCTGAGTCT 3360  
3420 AGTAATTTGTAAGCATCTCTGAATTAATAAAGCAAGTATTAAGTATGAGTCTGAGTCT 3479  
3361 AATACAGATTTCTCTCATATCTGATTTAGATTAATGAGAGAGAGAGAGAGAGAGAGAG 3420  
3480 AATACAGATTTCTCTCATATCTGATTTAGATTAATGAGAGAGAGAGAGAGAGAGAGAG 3539  
3421 CATGCTATGAGTGTGTTGTTGAGACACCTGATGAGTCTGATGATGATGATGATGATGAT 3480  
3540 CATGCTATGAGTGTGTTGTTGAGACACCTGATGAGTCTGATGATGATGATGATGATGAT 3599  
3481 GAAGATACTAGTGTGTTGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3540  
3600 GAAGATACTAGTGTGTTGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3659  
3541 GTCCAGAAAGAGAGGCTTAGCAGAGAGTCTAGGCTTTACCCATACACATTTGGCTCAG 3600  
3660 GTCCAGAAAGAGAGGCTTAGCAGAGAGTCTAGGCTTTACCCATACACATTTGGCTCAG 3719  
3601 GGTTCAGGAAAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3660  
3720 GGTTCAGGAAAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3779  
3661 GAAGAGCTTCCGCTTCCAAACACTGTTATTTGTTAAAGTAAACATATACCTTCTCAG 3720  
3780 GAAGAGCTTCCGCTTCCAAACACTGTTATTTGTTAAAGTAAACATATACCTTCTCAG 3839  
3721 TCTACTAGCAGTACAGCCTGTTGCTACCGAGTGTCTCTTAAGAACACAGAGAGAAATTA 3780  
3840 TCTACTAGCAGTACAGCCTGTTGCTACCGAGTGTCTCTTAAGAACACAGAGAGAAATTA 3899  
3781 TTTATCTTAAGAAATGCTTAATTAAGTCTGAGTAAACAGAGTATATTGCAAGGAGATCT 3840  
3900 TTTATCTTAAGAAATGCTTAATTAAGTCTGAGTAAACAGAGTATATTGCAAGGAGATCT 3959  
3841 CAGGAACATCACCTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3900  
3960 CAGGAACATCACCTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4019  
3901 AGTGAATTTGGAAGACTTGAAGTCAATTAACAAACCCAGAGATCTTTGATGTTGTTCT 3960

```

Db 4020 AGTGAATTGGAGACTTGAAGTCAAAATACAAACACCCAGATCCTTTTGATTGGTTCT 4079
OY 3961 TCCAAACAATATGAGCATGAGTCTGAAGCCAGGAGTTGCTGATGATGACAAAGAAATTG 4020
Db 4080 TCCAAACAATATGAGCATGAGTCTGAAGCCAGGAGTTGCTGATGATGACAAAGAAATTG 4139
OY 4021 GTTTCAGATGATGAGAAAGAGAGAGGCGCTTGAAAGAAATATCAAGAAGAGCAAGC 4080
Db 4140 GTTTCAGATGATGAGAAAGAGAGAGGCGCTTGAAAGAAATATCAAGAAGAGCAAGC 4139
OY 4081 ATGATTCACACTTAGTGAAGAGCATCTGGGTGGAGGTGAAGTGAACCAACCGCTCTGAA 4140
Db 4200 ATGATTCACACTTAGTGAAGAGCATCTGGGTGGAGGTGAAGTGAACCAACCGCTCTGAA 4259
OY 4141 GACTGCTCAGGCGTATCTCTCAGAGTGAATTTTAACTACAGCAGAGAGGATACCATG 4200
Db 4260 GACTGCTCAGGCGTATCTCTCAGAGTGAATTTTAACTACAGCAGAGAGGATACCATG 4319
OY 4201 CAACATTAACCTGATTAAGCTCCAGCAGGAATGGCTGAACCTAGAGCTGTGTAGAACAG 4260
Db 4320 CAACATTAACCTGATTAAGCTCCAGCAGGAATGGCTGAACCTAGAGCTGTGTAGAACAG 4379
OY 4261 CATGGAGGCGAGCTTCTTAACAGCTACCTCCATCATATAGTGAAGTCTTCCCTTGAG 4320
Db 4380 CATGGAGGCGAGCTTCTTAACAGCTACCTCCATCATATAGTGAAGTCTTCCCTTGAG 4439
OY 4321 GACCTGCGAAATCCAGAACAAAGCACATCAGAAAAAGCAGTATTAACTTACAGAAAGT 4380
Db 4440 GACCTGCGAAATCCAGAACAAAGCACATCAGAAAAAGCAGTATTAACTTACAGAAAGT 4499
OY 4381 AGTGAATACCTTATAGCCAGAAATCCAGAAAGCCTTCTGCTGACAAAGTTTGAAGTGTCT 4440
Db 4500 AGTGAATACCTTATAGCCAGAAATCCAGAAAGCCTTCTGCTGACAAAGTTTGAAGTGTCT 4559
OY 4441 GCAGATAGTTTACCAAGTAAATAAAGAACAGAGAGTGAAGAGTCAATCCCTTTTAA 4500
Db 4560 GCAGATAGTTTACCAAGTAAATAAAGAACAGAGAGTGAAGAGTCAATCCCTTTTAA 4619
OY 4501 TGCCCATCATTAAGATAGTGTGATGACAGTGTCTGAGAGTCTTGAGATAGA 4560
Db 4620 TGCCCATCATTAAGATAGTGTGATGACAGTGTCTGAGAGTCTTGAGATAGA 4679
OY 4561 AACTACCATCTCAGAGAGAGTCAATTAAGTGTGATGATGAGAGAGTCAATCCCTTTTAA 4620
Db 4680 AACTACCATCTCAGAGAGAGTCAATTAAGTGTGATGATGAGAGAGTCAATCCCTTTTAA 4739
OY 4621 GAGTCTGGGCGACAGATTTTGAAGAAACATCTTACTTCCCAAGGCAAGATCTAGAGGA 4680
Db 4740 GAGTCTGGGCGACAGATTTTGAAGAAACATCTTACTTCCCAAGGCAAGATCTAGAGGA 4799
OY 4681 ACCCTTACCTGGAATCTGGAATCAGCCTTCTCTGATGACCTGGAATCTGATCCTTCT 4740
Db 4800 ACCCTTACCTGGAATCTGGAATCAGCCTTCTCTGATGACCTGGAATCTGATCCTTCT 4859
OY 4741 GAAGCAGAGGCGGAGTGTGATGAGTGTGTCACATCCATCTTCAACCTCTGATG 4800
Db 4860 GAAGCAGAGGCGGAGTGTGATGAGTGTGTCACATCCATCTTCAACCTCTGATG 4919
OY 4801 AAAGTTCCCAATTTGAAGTTGCAAGATCTGCCAGAGTCCAGCTGCTCTATACTACT 4860
Db 4920 AAAGTTCCCAATTTGAAGTTGCAAGATCTGCCAGAGTCCAGCTGCTCTATACTACT 4979
OY 4861 GATACCTGCTGGGTATATGCAATGGAAGAAAGTGTGAGCAGGAGGAGAGCCAGAAATTGACA 4920
Db 4980 GATACCTGCTGGGTATATGCAATGGAAGAAAGTGTGAGCAGGAGGAGAGCCAGAAATTGACA 5039
OY 4921 GCTTCAACAGAAAGGCTCAACAAAGAAATGTCATGAGTGTGTGAGGCTTGACCCAGAA 4980
Db 5040 GCTTCAACAGAAAGGCTCAACAAAGAAATGTCATGAGTGTGTGAGGCTTGACCCAGAA 5099
OY 4981 GAATTTATGCTGCTGATACAAATTTGCGAGAAACCCAGATCACTTAACTAATCAAT 5040
Db 5100 GAATTTATGCTGCTGATACAAATTTGCGAGAAACCCAGATCACTTAACTAATCAAT 5159

```

```

OY 5041 ACTGAGAGACTACTCATGTTTATGAAACAGATGCTGAGTTTGTGTGACGAGACA 5100
Db 5160 ACTGAGAGACTACTCATGTTTATGAAACAGATGCTGAGTTTGTGTGACGAGACA 5219
OY 5101 CTGAAATATTTTCTAGGAATGGGAGAGGAAATGGGTAGTATTTCTGAGTACC 5160
Db 5220 CTGAAATATTTTCTAGGAATGGGAGAGGAAATGGGTAGTATTTCTGAGTACC 5279
OY 5161 CAGCTATTTAAGAAAGAAATGCTGATGAGCATGATTTTGAAGTCAAGAGAGATGTG 5220
Db 5280 CAGCTATTTAAGAAAGAAATGCTGATGAGCATGATTTTGAAGTCAAGAGAGATGTG 5339
OY 5221 GTCAATGAGAAAGAACCCAGAGTCCAAAGCAGAGAGAGATCCAGAGAGAAAGATC 5280
Db 5340 GTCAATGAGAAAGAACCCAGAGTCCAAAGCAGAGAGAGATCCAGAGAGAAAGATC 5399
OY 5281 TTCAGGGGCTAGAAATCTGTTGCTATGAGGCGCTTACCAACATGCCAGATCAACTG 5340
Db 5400 TTCAGGGGCTAGAAATCTGTTGCTATGAGGCGCTTACCAACATGCCAGATCAACTG 5459
OY 5341 GAATGAGTGTACAGCTGTGTGCTCTCTGATGAGAGAGTCTTATCACTTACCTT 5400
Db 5460 GAATGAGTGTACAGCTGTGTGCTCTCTGATGAGAGAGTCTTATCACTTACCTT 5519
OY 5401 GGCACAGTGTCCACCAATTTGTTGTGACAGCCAGATGCTGAGACAGAGACAAATGAC 5460
Db 5520 GGCACAGTGTCCACCAATTTGTTGTGACAGCCAGATGCTGAGACAGAGACAAATGAC 5579
OY 5461 TTCATGCAATTTGGCAGATGTGTAGAGCACCCTGTGTGACCCAGATGGGTGTGGAC 5520
Db 5580 TTCATGCAATTTGGCAGATGTGTAGAGCACCCTGTGTGACCCAGATGGGTGTGGAC 5639
OY 5521 AGGTAGACACTTACAGTCCAGAGTGGACACTACCTGATATCCCAATCCCCAC 5580
Db 5640 AGGTAGACACTTACAGTCCAGAGTGGACACTACCTGATATCCCAATCCCCAC 5699
OY 5581 AGCCACTAC 5589
Db 5700 AGCCACTAC 5708

RESULT 2
AAC60795
ID AAC60795 standard; cDNA: 5711 BP.
XX
AC AAC60795;
XX
DT 07-FEB-2001 (first entry)
XX
DE Human BRCA1 (om13) nucleotide sequence SEQ ID NO:5.
XX
KW Human: BRCA1: chromosome 17: 17q21; breast cancer; ovarian cancer;
KW gene therapy; diagnosis; cytostatic; genetic susceptibility; mutation;
KW polymorphism; identification; ss.
XX
OS Homo sapiens.
XX
PN US6130322-A.
PD 10-OCT-2000.
XX
PF 06-MAY-1998; 98US-0074476.
XX
PR 12-FEB-1996; 96US-0598591.
PR 12-DEC-1997; 97US-0798691.
XX
PA (GENE-) GENE LOGIC INC.
PI Zeng B, Thurber D, Olson SJ, Alvares CP, Allen ACP, Murphy PD;
PI Critz BS;
XX
DR WPI: 2000-646756/62.

```





1681	TCCTATTCGAATGAGAAAAATCCTAACCCATPAGAAATCAGTCGAAAAAGAAATCGCTTTC	174
1800	TCCTATTCGAATGAGAAAAATCCTAACCCATPAGAAATCAGTCGAAAAAGAAATCGCTTTC	1859
1741	AAAAAGAAAGCTGAACTCTATAAGCAGCAGTATTAAGCAATATGGAACCTCGAATTTAAATATC	1800
1860	AAAAAGAAAGCTGAACTCTATAAGCAGCAGTATTAAGCAATATGGAACCTCGAATTTAAATATC	1919
1801	CACAAATTCAAAAGCACCTTAAAAAGAAATAGGCTGAGAGCAAGTCTTCCACGACATATTT	1860
1920	CACAAATTCAAAAGCACCTTAAAAAGAAATAGGCTGAGAGCAAGTCTTCCACGACATATTT	1979
1861	CATGGCTGTGAACCTGATGTGACGTGAATATTAAGCCCACTTAATTTGATGCAATTTGCA	1920
1980	CATGGCTGTGAACCTGATGTGACGTGAATATTAAGCCCACTTAATTTGATGCAATTTGCA	2039
1921	ATTGATAGTTGTTTACGACGTGAAGAGATTAAGAAAAAAAGTACAAACCAATGCCAGTTC	1980
2040	ATTGATAGTTGTTTACGACGTGAAGAGATTAAGAAAAAAAGTACAAACCAATGCCAGTTC	2099
1981	AGGCACAGCAGAAACCTTACAACTCATGGAAGTTAAAGAACCTGCAACTGGAGCCACAAG	2040
2100	AGGCACAGCAGAAACCTTACAACTCATGGAAGTTAAAGAACCTGCAACTGGAGCCACAAG	2159
2041	AGTACAGCCCAAAAGSACACAGACAGTAAAGACATGACATGATACTTTCGCCAGAGCTG	2100
2160	AGTACAGCCCAAAAGSACACAGACAGTAAAGACATGACACCGATCTTCCACAGAGCTG	2219
2101	AAGTTAAACAATGACACCTGTTCTTTTACTAAGTTTCAATTAACACAGTAAAGAA	2160
2220	AAGTTAAACAATGACACCTGTTCTTTTACTAAGTTTCAATTAACACAGTAAAGAA	2279
2161	TTTGTGCAATCCTAGCCCTTCCAGAGAGAGAAAAAGAGAAACTAGAAACAGTTAAAGTG	2220
2280	TTTGTGCAATCCTAGCCCTTCCAGAGAGAGAAAAAGAGAAACTAGAAACAGTTAAAGTG	2339
2221	TCTAATATATGCTGAAGAGACCCCAAGAGATCTCATGTTAACTGAGAAAGGGTTTCCAACT	2280
2340	TCTAATATATGCTGAAGAGACCCCAAGAGATCTCATGTTAACTGAGAAAGGGTTTCCAACT	2399
2281	GAAAGATCTGTAGAGAGTAGCAGTAAATTCATTTGGTACTGCTAGTATATGGCACTCAG	2340
2400	GAAAGATCTGTAGAGAGTAGCAGTATTTTCATTTGGTACTGCTAGTATATGGCACTCAG	2459
2341	GAAAGATCTGTTACTGTGGAAGTTAGCAGCTGTAGGGAAGGCCAAAAACAGAACCAATPAA	2400
2460	GAAAGATCTGTTACTGTGGAAGTTAGCAGCTGTAGGGAAGGCCAAAAACAGAACCAATPAA	2519
2401	TGTGTGATCAGTGTGTGAGCAGATTTTGAAAAACCCCAAGGACACTAATTCATGGTGTTCCAA	2460
2520	TGTGTGATCAGTGTGTGAGCAGATTTTGAAAAACCCCAAGGACACTAATTCATGGTGTTCCAA	2579
2461	GATAATATGAATGACACAGAGGCTTTTAAATATCCCATGCGGACATGAAGTTAACACAGT	2520
2580	GATAATATGAATGACACAGAGGCTTTTAAATATCCCATGCGGACATGAAGTTAACACAGT	2639
2521	CGGGAAACAGACATAGAAATGGAAGAAAGTAACTGATGCTCAGATTTTGACAGAAATACA	2580
2640	CGGGAAACAGACATAGAAATGGAAGAAAGTAACTGATGCTCAGATTTTGACAGAAATACA	2699
2581	TTTCAAGGTTTCAAAAGCGCCAGTCAATTTGCTGTGTTTCAATTCAGAGAAATGCAAGAG	2640
2700	TTTCAAGGTTTCAAAAGCGCCAGTCAATTTGCTGTGTTTCAATTCAGAGAAATGCAAGAG	2759
2641	GAATGTGCAACATTTCTGTGCCACTGTGGGTGCTTTAAAGAAACAAAGTCCAAAAGTCACT	2700
2760	GAATGTGCAACATTTCTGTGCCACTGTGGGTGCTTTAAAGAAACAAAGTCCAAAAGTCACT	2819
2701	TTTGAATGTGAAAAAGAAAGAAATCAAGAGAAAGATGATCTAATATATCAAGCCGTGA	2760
2820	TTTGAATGTGAAAAAGAAAGAAATCAAGAGAAAGATGATCTAATATATCAAGCCGTGA	2879
2761	CAGACAGTTAATATCACTCAGAGCTTTCCTGTGCTGAGTGCAGAAAGATTAAGCCAGTTGAT	2820

Db	2880	CAGACAGTTAATATACCTACGAGGCTTCTCTGTGGTTGGTCAGAAAGATTAAGCCAGTTAT	2933
QY	2821	AATGGCCAATGTAGTATCAAAAGAGGCGCTAGGTTTTGTATATCATCTCAGTTTCAGAGGC	2880
Db	2940	AATGCCAAATGTAGTATATCAAAAGAGGCGCTAGTTTGTGCTATCATCTCAGTTTCAGAGGC	2999
QY	2881	AACGAAATGGACCTATTAATCTCCAAATTAACATGACCTTTTTCAAAAACCCATATCGTATA	2940
Db	3000	AACGAAATGGACCTATTAATCTCCAAATTAACATGACCTTTTTCAAAAACCCATATCGTATA	3055
QY	2941	CCACACATTTTCCCATCAGCTCATTTGTGTTAAACTTAATATGTAACAAAATCTGCTAAG	3000
Db	3060	CCACACATTTTCCCATCAGCTCATTTGTGTTAAACTTAATATGTAACAAAATCTGCTAAG	3119
QY	3001	GAAACTTTGAGGAACATTCATATGTCACCTGAAAGAGAATGGAAATGAGAACATTTCCA	3060
Db	3120	GAAACTTTGAGGAACATTCATATGTCACCTGAAAGAGAATGGAAATGAGAACATTTCCA	3179
QY	3061	AGTACAGTGAAGCAATTAAGCCGTATATTAACATTAGAGAAATGTTTTTAAAGAACCCAGC	3120
Db	3180	AGTACAGTGAAGCAATTAAGCCGTATATTAACATTAGAGAAATGTTTTTAAAGAACCCAGC	3233
QY	3121	TCACCAATATTAATGAAGTAGGCTCCGCTACTATGTAAGTGGGCTCCAGATTAATGA	3180
Db	3240	TCACCAATATTAATGAAGTAGGCTCCGCTACTATGTAAGTGGGCTCCAGATTAATGA	3299
QY	3181	ATAGGTTCCAGTGAAGAAACATTCACAGACAGAACAGGTAGAACACAGAGGCCAAAAATG	3240
Db	3300	ATAGGTTCCAGTGAAGAAACATTCACAGACAGAACAGGTAGAACACAGAGGCCAAAAATG	3355
QY	3241	AATGCTATGCTTACATTAAGGGTTTTGCAACCTGAGGCTATTAACAAAGCTTCTCTGGA	3300
Db	3360	AATGCTATGCTTACATTAAGGGTTTTGCAACCTGAGGCTATTAACAAAGCTTCTCTCTGGA	3419
QY	3301	AGTATTTGAAGCATCTGGAATTAAGGAACAGAAATGTAAGAAATGTTTCAGACTGT	3360
Db	3420	AGTATTTGAAGCATCTGGAATTAAGGAACAGAAATGTAAGAAATGTTTCAGACTGT	3479
QY	3361	AATACAGATTTCTCTCATATCTGATTTAGATTAACCTTAGAACAGCCATGGAAGTATAGT	3420
Db	3480	AATACAGATTTCTCTCATATCTGATTTAGATTAACCTTAGAACAGCCATGGAAGTATAGT	3539
QY	3421	CATGATCTCAGGTTTGTTCGACACACCTGATGACCTGTTAGATGATGTAAGTAAG	3480
Db	3540	CATGATCTCAGGTTTGTTCGACACACCTGATGACCTGTTAGATGATGTAAGTAAG	3599
QY	3481	GAAATATCTAGTTTGTGTAAGAAATACATTAAGAAATCTGCTGTTTAGCAAAAGC	3540
Db	3600	GAAATATCTAGTTTGTGTAAGAAATACATTAAGAAATCTGCTGTTTAGCAAAAGC	3659
QY	3541	GTCCGAAAGAGAGACTTAGCAGAGAGTCTATAGCCCTTTCACCCATATACATTTGGCTCAG	3600
Db	3660	GTCCGAAAGAGAGACTTAGCAGAGAGTCTATAGCCCTTTCACCCATATACATTTGGCTCAG	3719
QY	3601	GCTTACCGAAGAGGGGCCAAGAATTTAGAGTCTCTAGAGAGAACTTATCTAGTAGAGAT	3660
Db	3720	GCTTACCGAAGAGGGGCCAAGAATTTAGAGTCTCTAGAGAGAACTTATCTAGTAGAGAT	3779
QY	3661	GAAAGAGCTTCCCTGCTTCCACACCTTGTATTTGGTAAAGTAACATATACCTTCTCAG	3720
Db	3780	GAAAGAGCTTCCCTGCTTCCACACCTTGTATTTGGTAAAGTAACATATACCTTCTCAG	3839
QY	3721	TCTACTAGAGCTTAGACAGCTTGTATACGAGAGTGTGCTTAGAGACACAGAGAGAAATTTA	3780
Db	3840	TCTACTAGAGCTTAGACAGCTTGTATACGAGAGTGTGCTTAGAGACACAGAGAGAAATTTA	3899
QY	3781	TTATCATTTGAAGAAATAGCTTAAATGACTGACGTAAACAGGTAATATTTGGCAAGGACATCT	3840
Db	3900	TTATCATTTGAAGAAATAGCTTAAATGACTGACGTAAACAGGTAATATTTGGCAAGGACATCT	3959
QY	3841	CAGGAACATCACCTTAGTAGAGAAACAAATGTTCTGCTAGCTGTTTTCTTCACAGTGC	3900

Db 3960 CAGGACATCACTTACTGAGGAAACAAATGTTCTGCTAGCTGTTGTTCTTCACAGTGC 4019  
 QY 3901 AGTAATTTGAGAACTGCTGCTGCAAAATACAAACCCAGAGATCCTTCTTGATGGTCT 3960  
 Db 4020 AGTAATTTGAGAACTGCTGCTGCAAAATACAAACCCAGAGATCCTTCTTGATGGTCT 4079  
 QY 3961 TCCAAACAAATGAGGCACTGCTGCAAAAGCCAGGAGTGGTCTGAGTGCACAAAGATG 4020  
 Db 4080 TCCAAACAAATGAGGCACTGCTGCAAAAGCCAGGAGTGGTCTGAGTGCACAAAGATG 4139  
 QY 4021 GTTTCAGATATGAGAAAGAGAGAGGCGCTTGGAGAGAAATATCAAGAGAGCAAGC 4080  
 Db 4140 GTTTCAGATATGAGAAAGAGAGAGGCGCTTGGAGAGAAATATCAAGAGAGCAAGC 4199  
 QY 4081 ATGATTTCAAACTTAGGTGAGAGAGCATGCGGGTGTGAGAGTGAACAGCGTCTCGAA 4140  
 Db 4200 ATGATTTCAAACTTAGGTGAGAGAGCATGCGGGTGTGAGAGTGAACAGCGTCTCGAA 4259  
 QY 4141 GACTGCTCAGGGGCTATCCTCTCAGAGTGCATTTTAACTGACAGCAGAGGAGATACCATG 4200  
 Db 4260 GACTGCTCAGGGGCTATCCTCTCAGAGTGCATTTTAACTGACAGCAGAGGAGATACCATG 4319  
 QY 4201 CAACATTAACCTGATTAAGCTCCAGCAGAAATGGCTGAACTAGAACTGTTTGAACAG 4260  
 Db 4320 CAACATTAACCTGATTAAGCTCCAGCAGAAATGGCTGAACTAGAACTGTTTGAACAG 4379  
 QY 4261 CATGGAGGAGGCTTCTTAACAGTACCTTCCATCATTAAGTGAAGTGAAGTGAAGTGAAG 4320  
 Db 4380 CATGGAGGAGGCTTCTTAACAGTACCTTCCATCATTAAGTGAAGTGAAGTGAAGTGAAG 4439  
 QY 4321 GACCTGGGAAATCCAGAAACAAAGCACAATCAGAAAAAGCAGTATTTACTTCACAGAAAGT 4380  
 Db 4440 GACCTGGGAAATCCAGAAACAAAGCACAATCAGAAAAAGCAGTATTTACTTCACAGAAAGT 4499  
 QY 4381 AGTGAATACCTTAAAGCCAGAAATCCAGAAAGGCTTCTCTGCAAGAGTTTGAGTCT 4440  
 Db 4500 AGTGAATACCTTAAAGCCAGAAATCCAGAAAGGCTTCTCTGCAAGAGTTTGAGTCT 4559  
 QY 4441 GCAGATAGTTCTACCACTAAATAATTAAGAACAGAGAGTGGAAAGGATCCCTCTTAAA 4500  
 Db 4560 GCAGATAGTTCTACCACTAAATAATTAAGAACAGAGAGTGGAAAGGATCCCTCTTAAA 4619  
 QY 4501 TGGCCATCATTTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4560  
 Db 4620 TGGCCATCATTTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4679  
 QY 4561 AACTACCATCTCAAGAGAGAGCTCATTAAGTGTGTATGTGAGAGAGCAACAGCTGGA 4620  
 Db 4680 AACTACCATCTCAAGAGAGAGCTCATTAAGTGTGTATGTGAGAGAGCAACAGCTGGA 4739  
 QY 4621 GAGCTGGGCGACACGATTTGAGCGAAACATCTTACTTGGCAAGGCAAGATCTAGAGGA 4680  
 Db 4740 GAGCTGGGCGACACGATTTGAGCGAAACATCTTACTTGGCAAGGCAAGATCTAGAGGA 4799  
 QY 4681 ACCCTTACCTGAAATCTGAAATCAGCCTTCTCTGATGATGATGATGATGATGATGATG 4740  
 Db 4800 ACCCTTACCTGAAATCTGAAATCAGCCTTCTCTGATGATGATGATGATGATGATGATG 4859  
 QY 4741 GAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4800  
 Db 4860 GAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4919  
 QY 4801 AAAGTTCCCAATTTGAAGTTGAGAAATGAGCAATGAGCAAGTGCAGTGCCTGCTACTACT 4860  
 Db 4920 AAAGTTCCCAATTTGAAGTTGAGCAATGAGCAATGAGCAAGTGCAGTGCCTGCTACTACT 4979  
 QY 4861 GATAGTCTGCTGATATATGCAATGAGAAAGTGTGAGAGAGAGAGAGAGAGAGAGAG 4920  
 Db 4980 GATAGTCTGCTGATATATGCAATGAGAAAGTGTGAGAGAGAGAGAGAGAGAGAGAG 5039  
 QY 4921 GCTTCAACAGAAAGGCTCAACAAAGAGATGCTATGCTGCTGCTGCTGCTGCTGCTG 4980  
 Db 5040 GCTTCAACAGAAAGGCTCAACAAAGAGATGCTATGCTGCTGCTGCTGCTGCTGCTGCTG 5099

QY 4981 GAATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5040  
 Db 5100 GAATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5159  
 QY 5041 ACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5100  
 Db 5160 ACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5219  
 QY 5101 CTGAAATATTTTCTAGAGAAATGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5160  
 Db 5220 CTGAAATATTTTCTAGAGAAATGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5279  
 QY 5161 CAGTCTATTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5220  
 Db 5280 CAGTCTATTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5339  
 QY 5221 GTCATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5280  
 Db 5340 GTCATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5399  
 QY 5281 TTGAGGGGGCTAGAAATCTGCTGATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 5340  
 Db 5400 TTGAGGGGGCTAGAAATCTGCTGATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 5459  
 QY 5341 GAATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5400  
 Db 5460 GAATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5519  
 QY 5401 GGCACAGAGTGTCCACCAATTTGTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5460  
 Db 5520 GGCACAGAGTGTCCACCAATTTGTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5579  
 QY 5461 TTCCATGCAATTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5520  
 Db 5580 TTCCATGCAATTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5639  
 QY 5521 AGTGTACACTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5580  
 Db 5640 AGTGTACACTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5699  
 QY 5581 AGCCACTAC 5589  
 Db 5700 AGCCACTAC 5708  
 RESULT 3  
 AAT17493  
 ID AAT17493 standard; cDNA; 5914' BP.  
 XX  
 AC AAT17493;  
 XX  
 DT 02-OCT-1996 (first entry)  
 XX  
 DE Mutated BRCA1 coding sequence from PM06.  
 XX  
 KW Cancer therapy: breast and ovarian cancer predisposing gene; immunogen;  
 KW antibody production; germline alteration; probe; lesion neoplasia; human;  
 KW gene therapy; protein replacement therapy; protein mimetic; BRCA1; ds.  
 OS Homo sapiens.  
 XX  
 FH Key location/Qualifiers.  
 FT CDS 120..5711  
 FT /\*tag= a  
 FT /product= BRCA1 mutant  
 FT mutation 2731  
 FT /\*tag= b  
 FT /note= "C to T mutation"  
 XX  
 PN W09605306-A2.  
 XX  
 PD 22-FEB-1996.

XX 11-AUG-1995; 95WO-US10202.  
 XX 07-JUN-1995; 95US-0483553.  
 PR 12-AUG-1994; 94US-0289221.  
 PR 02-SEP-1994; 94US-0300266.  
 PR 16-SEP-1994; 94US-0308104.  
 PR 29-NOV-1994; 94US-0348824.  
 PR 24-MAR-1995; 95US-0409305.  
 PR 07-JUN-1995; 95US-0480784.  
 XX (MRI-) MRIAD GENETICS INC.  
 PA (CANC-) CANCER INST.  
 PA (RECH-) CENT RECH DU CHUL.  
 XX Durocher F, Emi M, Nakamura Y, Simard J, Shattuck-Eidens DM;  
 PI WPI: 1996-139702/14.  
 DR P-PSDB; AAR81535.  
 XX New nucleic acid and polypeptide for mutant or polymorphic BRCA1  
 PT gene - for diagnosis and therapy of human breast and ovarian cancer  
 PT and for diagnosing pre-disposition to these cancers  
 XX Claim 1; 218pp; English.

XX AAT17439-r17453 and AAT17455-r17529 represent mutations of the human  
 CC breast and ovarian cancer predisposing gene (BRCA1) (for wild type see  
 CC AAT17438). Proteins encoded by these mutations (see AAR81483-R81497 and  
 CC AAR81499-R81546) can be used as immunogens for antibody production.  
 CC These mutant BRCA1 genes have at least 1 mutation or polymorphism in  
 CC comparison to the wild type sequence. By detecting a germline alteration  
 CC in the wild type BRCA1 gene, a predisposition for breast and ovarian  
 CC cancer can be diagnosed. In one method, BRCA1 mRNA isolated from a tissue  
 CC sample from a subject has a probe, corresponding to a fragment of the  
 CC wild type sequence (or an allele-specific probe for one of these  
 CC mutations), added to it. The conditions allow for hybridisation of the  
 CC probe to the mRNA, and any hybridisation which occurs is detected.  
 CC Alternatively the BRCA1 gene in the tissue sample is isolated, and a  
 CC shift in electrophoretic mobility of single stranded DNA from the sample  
 CC on a non-denaturing polyacrylamide gel indicates a mutation. These  
 CC methods of detection can also diagnose a lesion neoplasia associated with  
 CC the BRCA1 locus. The methods may be used in gene therapy, protein  
 CC replacement therapy and protein mimetics, and may be used to screen for  
 CC drugs in cancer therapy.

XX Sequence 5914 BP; 2006 A; 1155 C; 1316 G; 1437 T; 0 other:

Query Match 100.0%; Score 5587.4; DB 17; Length 5914;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 5588; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ATGATTTATCTGCTCTTCCGTGTAAGATGACAAATGCTTAATGCTATGCGAATA 60  
 DB 120 ATGGATTTATCTGCTCTTCCGTGTAAGATGACAAATGCTTAATGCTATGCGAATA 179  
 OY 61 ATCTTAGAGTGTCCATCTGTCTGAGTGTGATCAGAACCTGTCTCCACAAGTGTGAC 120  
 DB 180 ATCTTAGAGTGTCCATCTGTCTGAGTGTGATCAGAACCTGTCTCCACAAGTGTGAC 239  
 OY 121 CACATATTTTGCATATTTTGCATGCTGAAACTTCTCACCCAGAGAAAGGCGCTTCACAG 180  
 DB 240 CACATATTTTGCATATTTTGCATGCTGAAACTTCTCACCCAGAGAAAGGCGCTTCACAG 299  
 OY 181 TGTCTTATATTAAGATATATATACCAAGAGAGCTTCAGAAAGTACGAGATTAGT 240  
 DB 300 TGTCTTATATTAAGATATATATACCAAGAGAGCTTCAGAAAGTACGAGATTAGT 359  
 OY 241 CAACCTTGTGAAGAGCTATTTGAAATCATTTTGTGCTTTCAGCTTACACAGGTTTGAG 300  
 DB 360 CAACCTTGTGAAGAGCTATTTGAAATCATTTTGTGCTTTCAGCTTACACAGGTTTGAG 419  
 OY 301 TATGCAAAACGCTATTAATTTTGCACAAAAGAAATTAATCTCTCCGACATCTAAAGAT 360

DB 420 TATGCAAAACGCTATTAATTTTGCACAAAAGAAATTAATCTCTCCGACATCTAAAGAT 479  
 OY 361 GAAGTTTCTATCTATCCAAAGATATGGCTACAGAAACCGTCCCAAAAAGCTTTACAGAGT 420  
 DB 480 GAAGTTTCTATCTATCCAAAGATATGGCTACAGAAACCGTCCCAAAAAGCTTTACAGAGT 539  
 OY 421 GAACCCGAAATATCTCTTGTGAGAAACAGCTCAGTGTCAACTCTCAACTTGA 480  
 DB 540 GAACCCGAAATATCTCTTGTGAGAAACAGCTCAGTGTCAACTCTCAACTTGA 599  
 OY 481 ACTGTGAGAACTCTGAGCAAAAGCAGCGATTAACAACCTCAAAAGAGTGTCTACATT 540  
 DB 600 ACTGTGAGAACTCTGAGCAAAAGCAGCGATTAACAACCTCAAAAGAGTGTCTACATT 659  
 OY 541 GAATTTGGAGTCTGATCTCTGAGATACCGTTAATAGCAACTTATTTGCGAGTGGGA 600  
 DB 660 GAATTTGGAGTCTGATCTCTGAGATACCGTTAATAGCAACTTATTTGCGAGTGGGA 719  
 OY 601 GATCAAGAAATTTGTACAAATCAACCCCTCAAGAGACAGGAGTGAATCAGTTGGATTCT 660  
 DB 720 GATCAAGAAATTTGTACAAATCAACCCCTCAAGAGACAGGAGTGAATCAGTTGGATTCT 779  
 OY 661 GCAAAAAGGCTGCTGTGAATTTTCTGAGACGAGTGTAAACAATCTGAACATCATCA 720  
 DB 780 GCAAAAAGGCTGCTGTGAATTTTCTGAGACGAGTGTAAACAATCTGAACATCATCA 839  
 OY 721 CCCAGTAATTAATTTGAAACCCACTGGAAGCCGAGCTGAGAGGATCCAGAAAG 780  
 DB 840 CCCAGTAATTAATTTGAAACCCACTGGAAGCCGAGCTGAGAGGATCCAGAAAG 899  
 OY 781 TATCAGAGTATGTTCTGTTCAACTTGCATGTGAGAGCCATGTGGCAAAATACTATGCC 840  
 DB 900 TATCAGAGTATGTTCTGTTCAACTTGCATGTGAGAGCCATGTGGCAAAATACTATGCC 959  
 OY 841 AGCTCATTAACAGCATGAGAACAGCAGTTTATTACTACTAAAGACAGATGATAGCA 900  
 DB 960 AGCTCATTAACAGCATGAGAACAGCAGTTTATTACTACTAAAGACAGATGATAGCA 1019  
 OY 901 AAGCTGAATTTCTGTAATTAAGCAAAACAGCCGTGTGAGCAAGAGCCACATATACGA 960  
 DB 1020 AAGCTGAATTTCTGTAATTAAGCAAAACAGCCGTGTGAGCAAGAGCCACATATACGA 1079  
 OY 961 TGGGCTGGAAGTAAGGAAACATGTAATGATAGCGGAGCTCCAGCAGCAAGAAAAAGTA 1020  
 DB 1080 TGGGCTGGAAGTAAGGAAACATGTAATGATAGCGGAGCTCCAGCAGCAAGAAAAAGTA 1139  
 OY 1021 GATCTGAATGCTGATCCCTGTGTGAGAGAAAGATGGAATTAAGCAGAAATGCTCATGC 1080  
 DB 1140 GATCTGAATGCTGATCCCTGTGTGAGAGAAAGATGGAATTAAGCAGAAATGCTCATGC 1199  
 OY 1081 TCAGAGAACTCTAGATATCTAGATGATGTTCCCTGTGATTAACCTTAATAGCAGCTTCAG 1140  
 DB 1200 TCAGAGAACTCTAGATATCTAGATGATGTTCCCTGTGATTAACCTTAATAGCAGCTTCAG 1259  
 OY 1141 AAGGTTAATGAGTGTTCCTCAAGAGTGAATGATGTTAGTGTGATGAGTACATGAT 1200  
 DB 1260 AAGGTTAATGAGTGTTCCTCAAGAGTGAATGATGTTAGTGTGATGAGTACATGAT 1319  
 OY 1201 GGGGAGTCTGAATCAATATCCCAAAGTAGTGTATTTGAGAGCTTTAAAGAGAGTAGAT 1260  
 DB 1320 GGGGAGTCTGAATCAATATCCCAAAGTAGTGTATTTGAGAGCTTTAAAGAGAGTAGAT 1379  
 OY 1261 GAATATTTGTTGTTCTCAGAGAAATATAGCTATGAGGAGGAGTCTCATGAGGCTTTA 1320  
 DB 1380 GAATATTTGTTGTTCTCAGAGAAATATAGCTATGAGGAGGAGTCTCATGAGGCTTTA 1439  
 OY 1321 ATATGTAAGAAAGTGAAGAGTCTCACTCAAAATCAGTAGAGAGTAAATATTAAGACAAAAATA 1380  
 DB 1440 ATATGTAAGAAAGTGAAGAGTCTCACTCAAAATCAGTAGAGAGTAAATATTAAGACAAAAATA 1499  
 OY 1381 TTTGGGAAAACTATGCGAAGAGAGGAGCCCTCCCACTTAAGCAGTGAATGCAAAAT 1440

Db 1500 TTGGGAAAACCTATCGAGAGAGCAAGCCCTCCCAACTTAAGCATATGTAAGAAAT 1559  
QY 1441 CTAATTATAGGAGCAATTTGTTACTGAGCCACAGATTAATACAGAGGTCCTCCACAAAT 1500  
Db 1560 CTAATTATAGGAGCAATTTGTTACTGAGCCACAGATTAATACAGAGGTCCTCCACAAAT 1619  
QY 1501 AAATTAAAGCGTAAAGAGACCTACATGAGCCCTTCATCTCAGAGATTTTATCAAGAA 1560  
Db 1620 AATTAAACCGTAAAGAGACCTACATGAGCCCTTCATCTCAGAGATTTTATCAAGAA 1679  
QY 1561 GCAGATTTGGCAGTTCATAAAGACTCCTGAATGATTAATCAGGGAATCAACAAACGAG 1620  
Db 1680 GCAGATTTGGCAGTTCATAAAGACTCCTGAATGATTAATCAGGGAATCAACAAACGAG 1739  
QY 1621 CAGAAATGCTAGTATGTAATTAATTAATAGTGCATGAGAAATTAACAAAGGTGAT 1680  
Db 1740 CAGAAATGCTAGTATGTAATTAATTAATAGTGCATGAGAAATTAACAAAGGTGAT 1799  
QY 1681 TCTATTGAGAAATGAGAAAATCTTAACCCATAGATCACTGAAAAAGATCTGCTTC 1740  
Db 1800 TCTATTGAGAAATGAGAAAATCTTAACCCATAGATCACTGAAAAAGATCTGCTTC 1859  
QY 1741 AAAAGCAAAAGTGAACCTTAAGCAGCAGTATAGCAATAGGAACTCGAATTAATATC 1800  
Db 1860 AAAAGCAAAAGTGAACCTTAAGCAGCAGTATAGCAATAGGAACTCGAATTAATATC 1919  
QY 1801 CACAATTCAAAAGCCCTTAAGAAATAGGCTGAGAGAGAACTCTTACCAGGCATAT 1860  
Db 1920 CACAATTCAAAAGCCCTTAAGAAATAGGCTGAGAGAGAACTCTTACCAGGCATAT 1979  
QY 1861 CATGGCCTTGAACCTAGTACTAGTAAATCTAAGCCCACTAATTTGTAATTCGAA 1920  
Db 1980 CATGGCCTTGAACCTAGTACTAGTAAATCTAAGCCCACTAATTTGTAATTCGAA 2039  
QY 1921 AATTGATAGTGTCTAGCAGTAAGAGATTAAGAAAAAAGTACACCAATGCCAGTC 1980  
Db 2040 AATTGATAGTGTCTAGCAGTAAGAGATTAAGAAAAAAGTACACCAATGCCAGTC 2099  
QY 1981 AAGGACAGAGAAACCTCAACCTAGGAAAGTAAAGAACTGCAACTGAGCCAGGAAG 2040  
Db 2100 AAGGACAGAGAAACCTCAACCTAGGAAAGTAAAGAACTGCAACTGAGCCAGGAAG 2159  
QY 2041 AGTAACAAAGCCAAATGAAAGAGACAGTAAGTAAGACATGACAGTACTTTCCAGAGCTG 2100  
Db 2160 AGTAACAAAGCCAAATGAAAGAGACAGTAAGTAAGACATGACAGTACTTTCCAGAGCTG 2219  
QY 2101 AAGTTAACAAATGACACCTGCTTTTCTTACTAAGTGTCAATACCAGTAACCTAAAGAA 2160  
Db 2220 AAGTTAACAAATGACACCTGCTTTTCTTACTAAGTGTCAATACCAGTAACCTAAAGAA 2279  
QY 2161 TTGTGCAATCTAGGCTTCCAGAGAGAAAGAAAGAACTGAACAGTTAAAGTG 2220  
Db 2280 TTGTGCAATCTAGGCTTCCAGAGAGAAAGAAAGAAAGAACTGAACAGTTAAAGTG 2339  
QY 2221 TCTAATTAATGCTGAAGACCCCAAGATCTCATGTTAAGTGAGAAAGGTTTTGCAAACT 2280  
Db 2340 TCTAATTAATGCTGAAGACCCCAAGATCTCATGTTAAGTGAGAAAGGTTTTGCAAACT 2399  
QY 2281 GAAAGATCTGTAGAGAGTGAAGCTATTTGATGTAAGTCTGTAAGTATAGGCACTGAG 2340  
Db 2400 GAAAGATCTGTAGAGAGTGAAGCTATTTGATGTAAGTCTGTAAGTATAGGCACTGAG 2459  
QY 2341 GAAAGATCTGTAGTATGGAAGTGAAGTGAAGGAAAGGCAAAACAAACCAATTA 2400  
Db 2460 GAAAGATCTGTATCTGGAAGTGAAGTGAAGGAAAGGCAAAACAAACCAATTA 2519  
QY 2401 TGTGTGAGTCACTGTGACACATTTTGAAGAAACCCCAAGGAGTAAATTCATGTTTCCAAA 2460  
Db 2520 TGTGTGAGTCACTGTGACACATTTTGAAGAAACCCCAAGGAGTAAATTCATGTTTCCAAA 2579  
QY 2461 GATTAATGAATGACAGAGAGCTTTAAGTATCCATTGGACATGAAGTAAACCACT 2520  
Db 2580 GATTAATGAATGACAGAGAGCTTTAAGTATCCATTGGACATGAAGTAAACCACT 2639

QY 2521 CGGGAACAAGCATAGAAATGCAAGAAAGTGAACCTTGATGCTCAGTATTTGCAGATPACA 2580  
Db 2640 CGGGAACAAGCATAGAAATGCAAGAAAGTGAACCTTGATGCTCAGTATTTGCAGATPACA 2699  
QY 2581 TTCAAGTTTCAAAAGCGCCAGTCAATTTGCTGTGTTTTCAAATCCAGAGAAATGACAGAG 2640  
Db 2700 TTCAAGTTTCAAAAGCGCCAGTCAATTTGCTGTGTTTTCAAATCCAGAGAAATGACAGAG 2759  
QY 2641 GAATGTGCAACATTTCTGTGCTCCACTGTGGTCTTTAAGAAACAAAGTCCAAAGTCACT 2700  
Db 2760 GAATGTGCAACATTTCTGTGCTCCACTGTGGTCTTTAAGAAACAAAGTCCAAAGTCACT 2819  
QY 2701 TTTGAATGTGAACAAAGAGAGAAATTCAGAGAAACAAATGAGTCTAATTCAGACCTGTA 2760  
Db 2820 TTTGAATGTGAACAAAGAGAGAAATTCAGAGAAACAAATGAGTCTAATTCAGACCTGTA 2879  
QY 2761 CAGACATTAATATCACTGAGGCTTTCTGTGTTGTCAGAAAGATAGCCAGTTGAT 2820  
Db 2880 CAGACATTAATATCACTGAGGCTTTCTGTGTTGTCAGAAAGATAGCCAGTTGAT 2939  
QY 2821 AATGCCAATGTATGTCAAGAGAGGCTAGGTTTGTCTATCATCTCAGTTCAAGGC 2880  
Db 2940 AATGCCAATGTATGTCAAGAGAGGCTAGGTTTGTCTATCATCTCAGTTCAAGGC 2999  
QY 2881 AAGCAACGAGACTCATTTCTCCAAATTAACATGAGACTTTTACAAACCATATTCGTATA 2940  
Db 3000 AAGCAACGAGACTCATTTCTCCAAATTAACATGAGACTTTTACAAACCATATTCGTATA 3059  
QY 2941 CCACCACTTTTCCATCAAGTCAATTTGTTAAACATTAATGTAAAGAAATCTGCTAGAG 3000  
Db 3060 CCACCACTTTTCCATCAAGTCAATTTGTTAAACATTAATGTAAAGAAATCTGCTAGAG 3119  
QY 3001 GAAACTTTGAGAGACATTCATGTCACCTGAAAGAGAAATGGAATGGAACATTTCCA 3060  
Db 3120 GAAACTTTGAGAGACATTCATGTCACCTGAAAGAGAAATGGAATGGAACATTTCCA 3179  
QY 3061 AGTACAGTGAAGCAATTTAGGCGGTATTAACATTAAGAAATTTTAAAGAGGCAGC 3120  
Db 3180 AGTACAGTGAAGCAATTTAGGCGGTATTAACATTAAGAAATTTTAAAGAGGCAGC 3239  
QY 3121 TCAAGCAATATTAATGAAGTAGGTTCCAGTACTAATGAAGAGTGCGTCCAGTATTAATGA 3180  
Db 3240 TCAAGCAATATTAATGAAGTAGGTTCCAGTACTAATGAAGAGTGCGTCCAGTATTAATGA 3299  
QY 3181 ATAGGTTCCAGTATGAAGAAACATTTCAAGCAGAACTAGTGAAGAGAGAGGCCAAATTTG 3240  
Db 3300 ATAGGTTCCAGTATGAAGAAACATTTCAAGCAGAACTAGTGAAGAGAGAGGCCAAATTTG 3359  
QY 3241 AATGCTATGCTTATGATTAAGGAGTTTTCAGAACCTGAGAGTCTATTAACAAAGTCTTCCGGA 3300  
Db 3360 AATGCTATGCTTATGATTAAGGAGTTTTCAGAACCTGAGAGTCTATTAACAAAGTCTTCCGGA 3419  
QY 3301 AGTAATTTGAAGCATCTGAATTAAGAAAGCAAGAAATGAAGAAATGAGTTCAGACTGTT 3360  
Db 3420 AGTAATTTGAAGCATCTGAATTAAGAAAGCAAGAAATGAAGAAATGAGTTCAGACTGTT 3479  
QY 3361 AATACAGATTTCTCTCATATCTGATTTCAAGTAACTTGAACAGGCTATGGGAATGAT 3420  
Db 3480 AATACAGATTTCTCTCATATCTGATTTCAAGTAACTTGAACAGGCTATGGGAATGAT 3539  
QY 3421 CATGATCTCAGGTTTGTCTGTGACACCTGATGACCTGTAGATATAGTGAATTAAG 3480  
Db 3540 CATGATCTCAGGTTTGTCTGTGACACCTGATGACCTGTAGATATAGTGAATTAAG 3599  
QY 3481 GAAGATTAAGTATTTGCTGAATTAAGCAATTAAGGAAGTTCGCTTTTATGACAAAGC 3540  
Db 3600 GAAGATTAAGTATTTGCTGAATTAAGCAATTAAGGAAGTTCGCTTTTATGACAAAGC 3659  
QY 3541 GTCCAGAAAGAGAGGCTTGAAGAGAGTCTAGCCCTTTAACCATTAACATTTGGCTCAG 3600  
Db 3660 GTCCAGAAAGAGAGGCTTGAAGAGAGTCTAGCCCTTTAACCATTAACATTTGGCTCAG 3719

QY	3601	GGTTACCCGAAGGGGGCCAAAGAAATTAAGTCCTCAGAAAGCAACTTATCTAGAGAGAT	3660
Db	3720	GGTATCCGAAGGGGGCCAAAGAAATTAAGTCCTCAGAAAGCAACTTATCTAGAGAGAT	3779
QY	3661	GAGAGAGCTTCCCTGGCTCCCAACACTTGTTATTTGGTAAAGTAAACAATATACCTTCTCAG	3720
Db	3780	GAAAGAGCTTCCCTGGCTCCCAACACTTGTTATTTGGTAAAGTAAACAATATACCTTCTCAG	3839
QY	3721	TCTACTAGGCATAGCAGCCGTTGCTACGAGTGTCTGTCTAAGAACACAGAGAGAAATTTA	3780
Db	3840	TCTACTAGGCATAGCAGCCGTTGCTACGAGTGTCTGTCTAAGAACACAGAGAGAAATTTA	3899
QY	3781	TTATCATTGANAATAAGCTTAATATCAGTCGAGTAAACAGGTAATATTTGGCAAGGCACT	3840
Db	3900	TTATCATTGANAATAAGCTTAATATCAGTCGAGTAAACAGGTAATATTTGGCAAGGCACT	3959
QY	3841	CAGAAACATCACCTTAGTAGAGAAACAAATGTTCTGCTACTGTTTCTTCTCAGAGTC	3900
Db	3960	CAGAAACATCACCTTAGTAGAGAAACAAATGTTCTGCTACTGTTTCTTCTCAGAGTC	4019
QY	3901	AGTGAATTTGGAAGACTTGACTGACAAATACAAACACCACGAGTCCCTTCTTGATTGTTCT	3960
Db	4020	AGTGAATTTGGAAGACTTGACTGACAAATACAAACACCACGAGTCCCTTCTTGATTGTTCT	4079
QY	3961	TCCAAACAAATAGGCATCAGTCGTAAAGCCAGAGAGTTGGTCTGAGTGAACAAGAAATTG	4020
Db	4080	TCCAAACAAATAGGCATCAGTCGTAAAGCCAGAGAGTTGGTCTGAGTGAACAAGAAATTG	4139
QY	4021	GTTTCAGATGATGAAGAAAGGAACGGGCTTGGAGAAATTAATCAAGAAGACCAAGC	4080
Db	4140	GTTTCAGATGATGAAGAAAGGAACGGGCTTGGAGAAATTAATCAAGAAGACCAAGC	4139
QY	4081	ATGGAATTAACCTTAGTGTAAGACAGCATCTGGGTGTGAGAGTGAACAACAGCGTCTGAA	4140
Db	4200	ATGGAATTAACCTTAGTGTAAGACAGCATCTGGGTGTGAGAGTGAACAACAGCGTCTGAA	4259
QY	4141	GACTGCTCAGGGCTTCCCTCTCAGAGTGAACATTTTAAACCACTGCACAGAGGATATCATG	4200
Db	4260	GACTGCTCAGGGCTTCCCTCTCAGAGTGAACATTTTAAACCACTGCACAGAGGATATCATG	4319
QY	4201	CAACATTAACCTGATTAAGCTCCAGACAGAAATGGCTGAACATAGAAAGCTGTGTAACAAG	4260
Db	4320	CAACATTAACCTGATTAAGCTCCAGACAGAAATGGCTGAACATAGAAAGCTGTGTAACAAG	4379
QY	4261	CATGGGAGCCAGCCTTCTTAACAGCTACCCCTTCATCATTAAGTGACTCTTGCCCTTGAG	4320
Db	4380	CATGGGAGCCAGCCTTCTTAACAGCTACCCCTTCATCATTAAGTGACTCTTGCCCTTGAG	4439
QY	4321	GACCTGGCAAAATCCAGAACAAAGCACAATCAAGAAAAAGAGATTAATACCTCACAAGAAAGT	4380
Db	4440	GACCTGGCAAAATCCAGAACAAAGCACAATCAAGAAAAAGAGATTAATACCTCACAAGAAAGT	4499
QY	4381	AGTGAATACCTTATATAGCCAGAAATCCAGAAAGCCTTCTGCTGCAAGTTTGAGTGCT	4440
Db	4500	AGTGAATACCTTATATAGCCAGAAATCCAGAAAGCCTTCTGCTGCAAGTTTGAGTGCT	4559
QY	4441	GCAGATAGTTCACAGTAATAAATAAAGAACACAGAGAGTGAAGAGTATCCCTTCTTAA	4500
Db	4560	GCAGATAGTTCACAGTAATAAATAAAGAACACAGAGAGTGAAGAGTATCCCTTCTTAA	4619
QY	4501	TGCCCATATTAAGATAGTGGTGCATGCACAAGTGTCTTGGGAGCTTTCAGAAATAGA	4560
Db	4620	TGCCCATATTAAGATAGTGGTGCATGCACAAGTGTCTTGGGAGCTTTCAGAAATAGA	4679
QY	4561	AACATACCATCTCAAGAGAGAGCTCATTAAGGTTGTTGATGTGGAGAGCAACACACTGAA	4620
Db	4680	AACATACCATCTCAAGAGAGAGCTCATTAAGGTTGTTGATGTGGAGAGCAACACACTGAA	4739
QY	4621	GAGTCTGGGACACAAGATTTGACGGAAACATCTTACTTGTGCCAAGCAAGATCTAGAGGA	4680
Db	4740	GAGTCTGGGACACAAGATTTGACGGAAACATCTTACTTGTGCCAAGCAAGATCTAGAGGA	4799
QY	4681	ACCCCTTAACCTGGAATCTGGAATACGCTCTTCTCTGATGACCCCTGAACTGATCTTCT	4740

Db	4800	ACCCCTTACCTGGAAATCTGGAAATCAGCCTTCTCTGATGACCCCTGAATCTGATCTTCT	4855
Qy	4741	GAAAGAGAGGCCCCAGAGTCAGCTCGTGTGGCAACATACATCTTCAACCTCTGCAATGG	4800
Db	4860	GAAAGACAGAGGCCCAAGAGTCAGCTCGTGTGGCAACATACATCTTCAACCTCTGCAATGG	4919
Qy	4801	AAAGTTCCTCCCAATTTGAAGTTCGAGAAATCTGCCAGAGTCCAGCTGCTGCTCATCTACT	4866
Db	4920	AAAGTTCCTCCCAATTTGAAGTTCGAGAAATCTGCCAGAGTCCAGCTGCTGCTCATCTACT	4979
Qy	4861	GATACGTGCGGGGTATTAATGCAATGGAAGAAATGTGAGCAGGAGAGAACCCGAATTGACA	4920
Db	4980	GATACGTGCGGGGTATTAATGCAATGGAAGAAATGTGAGCAGGAGAGAACCCGAATTGACA	5039
Qy	4921	GCTTCAACGAAAGGGTCAACAAAAGAAATGTCAATGCTGATGTGCTGCGCTGACCCAGAA	4980
Db	5040	GCTTCAACGAAAGGGTCAACAAAAGAAATGTCAATGCTGATGTGCTGCGCTGACCCAGAA	5099
Qy	4981	GAAATTTATGCTCGTGTACAGTTGTCGAGAAAACACCAATCATTAACTAATCTAAT	5040
Db	5100	GAAATTTATGCTCGTGTACAGTTGTCGAGAAAACACCAATCATTAACTAATCTAAT	5159
Qy	5041	ACTGAAGAGACATCACTAGTGTGTTGTAAGAAACAGATGCTGAGTTGTGTGTAAGGACA	5100
Db	5160	ACTGAAGAGACATCACTAGTGTGTTGTAAGAAACAGATGCTGAGTTGTGTGTAAGGACA	5219
Qy	5101	CTGAATATATTTTCTAGGAATTCGCGAGAGAAATGGAGTAGTACATATTTCTGCGTGACC	5160
Db	5220	CTGAATATATTTTCTAGGAATTCGCGAGAGAAATGGAGTAGTACATATTTCTGCGTGACC	5279
Qy	5161	CAGCTCTATTAAAGAAAGAAAATGCTGAATGACATGATTTTGAAGTCAGAGAGATGTG	5220
Db	5280	CAGCTCTATTAAAGAAAGAAAATGCTGAATGACATGATTTTGAAGTCAGAGAGATGTG	5339
Qy	5221	GTCATATGGAAGAACCCACAGGTCCAAAGGAGCAAGAAATCCAGACAGAAAGATC	5280
Db	5340	GTCATATGGAAGAACCCACAGGTCCAAAGGAGCAAGAAATCCAGACAGAAAGATC	5399
Qy	5281	TTTCAGGGGGCTGGAATATCTGTGCTATGGGCCCTTACCAACATGCCCACAGATCACTG	5340
Db	5400	TTTCAGGGGGCTGGAATATCTGTGCTATGGGCCCTTACCAACATGCCCACAGATCACTG	5459
Qy	5341	GAATGATGTGTAACAGCTGTGTGTGCTCTCTGTGTGTAAGAGACCTTTCATTCACCCCT	5400
Db	5460	GAATGATGTGTAACAGCTGTGTGTGCTCTCTGTGTGTAAGAGACCTTTCATTCACCCCT	5519
Qy	5401	GGCACAAGGTGTCCACCACATTTGTTGTGTGTCAGACAGATGCTGTGACAGAGACAAATGGC	5460
Db	5520	GGCACAAGGTGTCCACCACATTTGTTGTGTGTCAGACAGATGCTGTGACAGAGACAAATGGC	5579
Qy	5461	TTTCCATGCAATTTGGCAGATGTGTGATGACACCTGTGTGATGACCCGAGATGGGTGTGGAC	5520
Db	5580	TTTCCATGCAATTTGGCAGATGTGTGATGACACCTGTGTGATGACCCGAGATGGGTGTGGAC	5639
Qy	5521	AGTGTAGCACTCTTACAGTGTCCAGAGAGCTGTGACACTGATACCCAGATCCCCAC	5580
Db	5640	AGTGTAGCACTCTTACAGTGTCCAGAGAGCTGTGACACTGATACCCAGATCCCCAC	5699
Qy	5581	AGCCACTAC 5589	
Db	5700	AGCCACTAC 5708	
RESULT 4			
AAAT17495			
AAAT17495 standard: cDNA: 5914 BP.			
AAAT17495:			
02-OCT-1996 (first entry)			
Mutated BRCA1 coding sequence from PM09.			

XX Cancer therapy; breast and ovarian cancer predisposing gene; immunogen;  
 KW antibody production; germline alteration; probe; lesion neoplasia; human;  
 KM gene therapy; protein replacement therapy; protein mimetic; BRCA1; ds.  
 XX  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT CDS 120..5711  
 FT /tag- a  
 FT /product- BRCA1 mutant  
 FT mutation 2201  
 FT /\*tag- b  
 FT /note- "C to T mutation"  
 XX  
 PD WO9605306-A2.  
 XX  
 PD 22-FEB-1996.  
 XX  
 PF 11-AUG-1995; 95MO-US10202.  
 XX  
 PR 07-JUN-1995; 95US-0483553.  
 PR 12-AUG-1994; 94US-0289221.  
 PR 02-SEP-1994; 94US-0300266.  
 PR 16-SEP-1994; 94US-0308104.  
 PR 29-NOV-1994; 94US-0348824.  
 PR 24-MAR-1995; 95US-0409305.  
 PR 07-JUN-1995; 95US-0480784.  
 XX  
 PA (MYRI-) MYRIAD GENETICS INC.  
 PA (CANC-) CANCER INST.  
 PA (RECH-) CENT RECH DU CHUL.  
 XX  
 PI Durocher F, Emi M, Nakamura Y, Simard J, Shattuck-Eidens DM;  
 DR WPI; 1996-139702/14.  
 DR P-PSDB; AAR81481.  
 XX  
 PT New nucleic acid and polypeptide for mutant or polymorphic BRCA1  
 PT gene - for diagnosis and therapy of human breast and ovarian cancer  
 PT and for diagnosing pre-disposition to these cancers  
 PS  
 PS Claim 1; 218bp; English.  
 XX  
 CC AAT17439-T17453 and AAT17455-T17529 represent mutations of the human  
 CC breast and ovarian cancer predisposing gene (BRCA1) (for wild type see  
 CC AAT17438). Proteins encoded by these mutations (see AAR81483-R81497 and  
 CC AAR81499-R81546) can be used as immunogens for antibody production.  
 CC These mutant BRCA1 genes have at least 1 mutation or polymorphism in  
 CC comparison to the wild type sequence. By detecting a germline alteration  
 CC in the wild type BRCA1 gene, a predisposition for breast and ovarian  
 CC cancer can be diagnosed. In one method, BRCA1 mRNA isolated from a tissue  
 CC sample from a subject has a probe, corresponding to a fragment of the  
 CC wild type sequence (or an allele-specific probe for one of these  
 CC mutations), added to it. The conditions allow for hybridisation of the  
 CC probe to the mRNA, and any hybridisation which occurs is detected.  
 CC Alternatively the BRCA1 gene in the tissue sample is isolated, and a  
 CC shift in electrophoretic mobility of single stranded DNA from the sample  
 CC on a non-denaturing polyacrylamide gel indicates a mutation. These  
 CC methods of detection can also diagnose a lesion neoplasia associated with  
 CC the BRCA1 locus. The methods may be used in gene therapy, protein  
 CC replacement therapy and protein mimetics, and may be used to screen for  
 CC drugs in cancer therapy.  
 XX  
 SO Sequence 5914 BP; 2006 A; 1155 C; 1316 G; 1437 T; 0 other;

Query Match 100.0%; Score 5587.4; DB 17; Length 5914;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 5588; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AATGATTATTCCTGCTTGGCGTTGAAGAAGTACAAATGCTATATGCTATGCGAGAA 60  
 DB 120 AATGATTATTCCTGCTTGGCGTTGAAGAAGTACAAATGCTATATGCTATGCGAGAA 179

QY 61 ATCTTAGAGTGTCCCATCTGCTGAGTGTATCAAGGAACCTGTCTCCACAAAGTGTAC 120  
 DB 180 ATCTTAGAGTGTCCCATCTGCTGAGTGTATCAAGGAACCTGTCTCCACAAAGTGTAC 239  
 QY 121 CACATATTTTGGCAAAATTTTGCATGCTGAACCTTCTCAACAGAGAAAGGCTTCACAG 180  
 DB 240 CACATATTTTGGCAAAATTTTGCATGCTGAACCTTCTCAACAGAGAAAGGCTTCACAG 269  
 QY 181 TGTCTCTTATGTAAGATGTATTAACCAAAAGAGCTTACAAAGAAATAGAGATTAGT 240  
 DB 300 TGTCTCTTATGTAAGATGTATTAACCAAAAGAGCTTACAAAGAAATAGAGATTAGT 359  
 QY 241 CAACCTGTGAAGAGCTATTTGAATCATTTGTCTTTCAGCTTGACAGGTTTGGAG 300  
 DB 360 CAACCTGTGAAGAGCTATTTGAATCATTTGTCTTTCAGCTTGACAGGTTTGGAG 419  
 QY 301 TATGCAAAACAGCTATTAATTTTGGCAAAAGAAATTAACCTCTCAACATCTAAAGAT 360  
 DB 420 TATGCAAAACAGCTATTAATTTTGGCAAAAGAAATTAACCTCTCAACATCTAAAGAT 479  
 QY 361 GAAGTTTCTATCATCCAAAGTATGGGCTACAGAAACCGTCCCAAAAGCTTTCACAGT 420  
 DB 480 GAAGTTTCTATCATCCAAAGTATGGGCTACAGAAACCGTCCCAAAAGCTTTCACAGT 539  
 QY 421 GAACCCGAAATCTCTCTGCGAGAAACAGCTCTGATGCCAAGCTTCAACCTTGA 480  
 DB 540 GAACCCGAAATCTCTCTGCGAGAAACAGCTCTGATGCCAAGCTTCAACCTTGA 599  
 QY 481 ACTGTGAGAACTCTGAGGACAAAGCAGGATCAACCTCAAAAGAGCTGTCTACATT 540  
 DB 600 ACTGTGAGAACTCTGAGGACAAAGCAGGATCAACCTCAAAAGAGCTGTCTACATT 659  
 QY 541 GAATTGGGATCTGATTTCTTGAAGATACCGTTATTAAGCAACTTATTCAGTGTGGGA 600  
 DB 660 GAATTGGGATCTGATTTCTTGAAGATACCGTTATTAAGCAACTTATTCAGTGTGGGA 719  
 QY 601 GATCAGAAATTTGTACAAATCACCCTCAAGGAACCGGATGAATCAGTTGGATTCT 660  
 DB 720 GATCAGAAATTTGTACAAATCACCCTCAAGGAACCGGATGAATCAGTTGGATTCT 779  
 QY 661 GCAAAAAAGGCTCTGTGTAATTTTCTGAGAGGATGTAAACAATATGTAATCATCAA 720  
 DB 780 GCAAAAAAGGCTCTGTGTAATTTTCTGAGAGGATGTAAACAATATGTAATCATCAA 839  
 QY 721 CCCAGTATATATTTTGAACACCACTGAGAAACCGTGTACCTGAGAGGATCCAGAAAG 780  
 DB 840 CCCAGTATATATTTTGAACACCACTGAGAAACCGTGTACCTGAGAGGATCCAGAAAG 899  
 QY 781 TATCAGGGTGTCTGTTTCAAACTTGATGTGAGGATGTGACCAATATGATGATGCC 840  
 DB 900 TATCAGGGTGTCTGTTTCAAACTTGATGTGAGGATGTGACCAATATGATGATGCC 959  
 QY 841 AGCTCATTCAGAGTGAAGAACACAGCTTTATCTCACTAATAAGACAGATGATAGAA 900  
 DB 960 AGCTCATTCAGAGTGAAGAACACAGCTTTATCTCACTAATAAGACAGATGATAGAA 1019  
 QY 901 AAGGCTGAATCTGTATATAAAGCAACACCGTGTGTACCAAGAGGACCAATACAGA 960  
 DB 1020 AAGGCTGAATCTGTATATAAAGCAACACCGTGTGTACCAAGAGGACCAATACAGA 1079  
 QY 961 TGGGCTGGAAGTGAAGAAACATGATGATGAGCGGATCCAGACAGAAAAAGGTA 1020  
 DB 1080 TGGGCTGGAAGTGAAGAAACATGATGATGAGCGGATCCAGACAGAAAAAGGTA 1139  
 QY 1021 GATCTGAATGCTGATCCCTGTGTGAGAGAAAGAAATGAATAGCAGAAATGCCATGC 1080  
 DB 1140 GATCTGAATGCTGATCCCTGTGTGAGAGAAAGAAATGAATAGCAGAAATGCCATGC 1199  
 QY 1081 TCAGAGAAATCTTGAAGATGATGAGATGTTCTTGGATTAACATTAATAGCAGATTGAG 1140  
 DB 1200 TCAGAGAAATCTTGAAGATGATGAGATGTTCTTGGATTAACATTAATAGCAGATTGAG 1259



1141 AAAGTTAATGAGTGGTTTTCAGAAAGTATGAACCTGTTAGTCTGATGACTCAGAT 1200  
1260 AAAGTTAATGAGTGGTTTTCAGAAAGTATGAACCTGTTAGTCTGATGACTCAGAT 1319  
1201 GGGGAGTGTGATCAAAATGCCAAAGTATGATGTTGGAGCTTCTAAATGAGTAT 1260  
1320 GGGGAGTGTGATCAAAATGCCAAAGTATGATGTTGGAGCTTCTAAATGAGTAT 1379  
1261 GAATATTTCTGTTCTTCAGAGAAATAGACTTACTGGCCAGTATCTCATGAGCTTTA 1320  
1380 GAATATTTCTGTTCTTCAGAGAAATAGACTTACTGGCCAGTATCTCATGAGCTTTA 1439  
1321 ATATGTAAAGTGAAGAAGTCTACTCTCAATCAGTAGAGATATATTTGAGACAAATA 1380  
1440 ATATGTAAAGTGAAGAAGTCTACTCTCAATCAGTAGAGATATATTTGAGACAAATA 1499  
1381 TTTGGGAAACCTATCGGAAAGGCAAGCTCCCACTTAAGCATGATGATGAAAT 1440  
1500 TTTGGGAAACCTATCGGAAAGGCAAGCTCCCACTTAAGCATGATGATGAAAT 1559  
1441 CTAATTTATAGAGCAATTTGTTACTGAGCCAGATATATCAAGAGCGTCCCTCAGAAAT 1500  
1560 CTAATTTATAGAGCAATTTGTTACTGAGCCAGATATATCAAGAGCGTCCCTCAGAAAT 1619  
1501 AAATTTAAAGCGTAAAGAGACCTACATAGGCTTTCATCTGAGGATTTTATCAAGAA 1560  
1620 AAATTTAAAGCGTAAAGAGACCTACATAGGCTTTCATCTGAGGATTTTATCAAGAA 1679  
1561 GCAGATTTGGCAGTTCAAAAAGACTCTGTAATGATTAATCAGGGAATCAACCAAGGAG 1620  
1680 GCAGATTTGGCAGTTCAAAAAGACTCTGTAATGATTAATCAGGGAATCAACCAAGGAG 1739  
1621 CAGATGCTCAAGTATGATATTTACTATATAGTGTGATGAGATATTAACCAAGGAGTAT 1680  
1740 CAGATGCTCAAGTATGATATTTACTATATAGTGTGATGAGATATTAACCAAGGAGTAT 1799  
1681 TCTATTCAAGTATGAGAAAAATCTCTAACCATAAGATCTCTGAGAAAAAGATCTGTTTC 1740  
1800 TCTATTCAAGTATGAGAAAAATCTCTAACCATAAGATCTCTGAGAAAAAGATCTGTTTC 1859  
1741 AAAAGGAAGCTGAAGCTTATAGCAGCAGTATAGCAATATGGAATGGAATTAATATTC 1800  
1860 AAAAGGAAGCTGAAGCTTATAGCAGCAGTATAGCAATATGGAATGGAATTAATATTC 1919  
1801 CACAATTCAAAAGCACTTAAAAAGATAGCTGAGAGAGAGCTTCTTCAACAGGCAATAT 1860  
1920 CACAATTCAAAAGCACTTAAAAAGATAGCTGAGAGAGAGCTTCTTCAACAGGCAATAT 1979  
1861 CATGGCTTGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920  
1980 CATGGCTTGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2039  
1921 ATTGATATGTTGTTAGCAGTGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGA 1980  
2040 ATTGATATGTTGTTAGCAGTGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGA 2099  
1981 AGGCACAGCAGAAACCTCAACTCATGAGAGTAAAGAACTGCAACTGGAGCCAGAG 2040  
2100 AGGCACAGCAGAAACCTCAACTCATGAGAGTAAAGAACTGCAACTGGAGCCAGAG 2159  
2041 AGTAACAAGCAAAATGAACAGACAGTAAGAACATGACAGTATCTTTCCAGAGCTG 2100  
2160 AGTAACAAGCAAAATGAACAGACAGTAAGAACATGACAGTATCTTTCCAGAGCTG 2219  
2101 AAGTTAAACAAATGCACTGCTGTTCTTTTACTAGTGTCAAAATACAGTGAACCTTAAGAA 2160  
2220 AAGTTAAACAAATGCACTGCTGTTCTTTTACTAGTGTCAAAATACAGTGAACCTTAAGAA 2279  
2161 TTTGTCATCTAGCCTTCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220  
2280 TTTGTCATCTAGCCTTCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2339  
2221 TCTAATATGCTGAAGAGACCCCAAGATCTCATGTTAAGTGAGAAAGGTTTTCAGAACT 2280

2340 TCTAATATATGCTGAAGAGACCCCAAGATCTCATGTTAAGTGAGAAAGGTTTTCAGAACT 2399  
2281 GAAAGATCTGTGAGAGTACGAGTATTCATGATGATGATGATGATGATGATGATGATGAT 2340  
2400 GAAAGATCTGTGAGAGTACGAGTATTCATGATGATGATGATGATGATGATGATGATGAT 2459  
2341 GAAAGATCTGTGAGAGTACGAGTATTCATGATGATGATGATGATGATGATGATGATGAT 2400  
2460 GAAAGATCTGTGAGAGTACGAGTATTCATGATGATGATGATGATGATGATGATGATGAT 2519  
2401 TGTGTAGTACGAGTACGAGTATTCATGATGATGATGATGATGATGATGATGATGATGAT 2460  
2520 TGTGTAGTACGAGTACGAGTATTCATGATGATGATGATGATGATGATGATGATGATGAT 2579  
2461 GATTAATAGAAATGACACAGAGGCTTTAAGTATTCATGATGATGATGATGATGATGATGAT 2520  
2580 GATTAATAGAAATGACACAGAGGCTTTAAGTATTCATGATGATGATGATGATGATGATGAT 2639  
2521 CGGGAACCAAGCATAGAAATGGAAGAGTGAATGATGATGATGATGATGATGATGATGATGAT 2580  
2640 CGGGAACCAAGCATAGAAATGGAAGAGTGAATGATGATGATGATGATGATGATGATGATGAT 2699  
2581 TTTCAAGTTTCAAGGCGCAGTATTTGCTGTTTCAATCCAGAAATGCGAGAGAG 2640  
2700 TTTCAAGTTTCAAGGCGCAGTATTTGCTGTTTCAATCCAGAAATGCGAGAGAG 2759  
2641 GAATGTGCAACATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700  
2760 GAATGTGCAACATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2819  
2701 TTTGAATGTGAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2760  
2820 TTTGAATGTGAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2879  
2761 CAGACAGTTAATATCTACGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820  
2880 CAGACAGTTAATATCTACGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2939  
2821 AATGCCAATGTATGATCAAGAGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2880  
2940 AATGCCAATGTATGATCAAGAGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2999  
2881 AACGAAGCTGACTATTAATCAATTAACATGACATTTTACAAACCCATATCGTATA 2940  
3000 AACGAAGCTGACTATTAATCAATTAACATGACATTTTACAAACCCATATCGTATA 3059  
2941 CCACCACTTTTCCATCAAGTATTTGTTTAAACTAATGTAAGAAAAATCTGCTAGAG 3000  
3060 CCACCACTTTTCCATCAAGTATTTGTTTAAACTAATGTAAGAAAAATCTGCTAGAG 3119  
3001 GAAACCTTTGAGAGCAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3060  
3120 GAAACCTTTGAGAGCAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3179  
3061 ACTACAGTACGACATTAAGCCGTATTAACATTAAGAGAAATGTTTAAAGAGCCAGC 3120  
3180 ACTACAGTACGACATTAAGCCGTATTAACATTAAGAGAAATGTTTAAAGAGCCAGC 3239  
3121 TCAAGCAATATTAAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3180  
3240 TCAAGCAATATTAAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3299  
3181 ATAGGTTCCAGTATGAGAAACATTCACAGCAAACTAGTATGAAACAGAGGCCAAATTTG 3240  
3300 ATAGGTTCCAGTATGAGAAACATTCACAGCAAACTAGTATGAAACAGAGGCCAAATTTG 3359  
3241 AATGCTATGCTTATGAGGTTTGGAGTGTGCAACCTAGGCTTAAACAAAGCTTCTCGGA 3300  
3360 AATGCTATGCTTATGAGGTTTGGAGTGTGCAACCTAGGCTTAAACAAAGCTTCTCGGA 3419  
3301 AGTAATGTGAAGCATCTGGAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3360



Dd	3420	AGTAATTGAAACATCCTGGAAATAAAAAACAAGAATATGAAGAAATGATGTTTCCACTGTT	3479
Oy	3361	AATACAGATTTCTCTCATATCTGTATTCAGATATACTTAGAACAGCCATGGAGAGT	3420
Dd	3480	AATACAGATTTCTCTCATATCTGTATTCAGATATACTTAGAACAGCCATGGAGAGT	3539
Oy	3421	CATGATCTCAGGTTTTGTTCTGAGACACCCTATGACCTGTTAATGATGCGTAATTAAG	3480
Dd	3540	CATGATCTCAGGTTTTGTTCTGAGACACCCTATGATGATGATGATGCGTAATTAAG	3599
Oy	3481	GAGATTAATACTTTTGCGAAATAGACATTAAGGAAGTCTGCTGTTTTAGCAAAAGC	3540
Dd	3600	GAGATTAATACTTTTGCGAAATAGACATTAAGGAAGTCTGCTGTTTTAGCAAAAGC	3659
Oy	3541	GTCGAAAGAGAGAGCTTAGCAGAGAGTCTTAGGCCCTTACCCATACATTTGGCTCAG	3600
Dd	3660	GTCGAGAAAGAGAGAGCTTAGCAGAGAGTCTTAGGCCCTTACCCATACATTTGGCTCAG	3719
Oy	3601	GATTACCGAGAGAGGGGCCAAGAAATTAGAGTCTCAGAGAGAACTTATCTAGTGAAGAT	3660
Dd	3720	GATTACCGAGAGAGGGGCCAAGAAATTAGAGTCTCAGAGAGAACTTATCTAGTGAAGAT	3779
Oy	3661	GAAAGAGTCTCCCTGCTTCCAACTGTTATTTGGTAAAGTAAACAAATATCCTTCCAG	3720
Dd	3780	GAAAGAGTCTCCCTGCTTCCAACTGTTATTTGGTAAAGTAAACAAATATCCTTCCAG	3839
Oy	3721	TCTACTAGGCATPACAGCCGTTGCTACAGAGTGTGCTGTAGAACACAGAGAGAAATTA	3780
Dd	3840	TCTACTAGGCATPACAGCCGTTGCTACAGAGTGTGCTGTAGAACACAGAGAGAAATTA	3899
Oy	3781	TTATCATTTGAAGAAATAGCTTAAATGACTGACGTAAACAGGTAATATTGGCAAAGCATCT	3840
Dd	3900	TTATCATTTGAAGAAATAGCTTAAATGACTGACGTAAACAGGTAATATTGGCAAAGCATCT	3959
Oy	3841	CAGGAACATCACCTTAGTGAGAGAAAACAAATGTTCTGCTACTGTTTCTTCCACAGTGC	3900
Dd	3960	CAGGAACATCACCTTAGTGAGAGAAAACAAATGTTCTGCTACTGTTTCTTCCACAGTGC	4019
Oy	3901	AGTGAATTTGGAGAACTTGACTGCAAAATCAACAACCCAGAGATCCCTTCTGATTTGTTCT	3960
Dd	4020	AGTGAATTTGGAGAACTTGACTGCAAAATCAACAACCCAGAGATCCCTTCTGATTTGTTCT	4079
Oy	3961	TCCAACAATAGAGGCATCACTGTGAAGCCAGAGAGTTGGTCTGAGTGACAAAGAAATG	4020
Dd	4080	TCCAACAATAGAGGCATCACTGTGAAGCCAGAGAGTTGGTCTGAGTGACAAAGAAATG	4139
Oy	4021	GTTTCAGATGATGAAAGAAAGAGAAACGGGCTTGGAAAGAAATATCAAGAAAGACCAAGC	4080
Dd	4140	GTTTCAGATGATGAAAGAAAGAGAAACGGGCTTGGAAAGAAATATCAAGAAAGACCAAGC	4199
Oy	4081	ATGATTTTAAACTTAGGTGAAGCAGCATCTGGGTGAGAGTGAACCAAGCGTCTGAA	4140
Dd	4200	ATGATTTTAAACTTAGGTGAAGCAGCATCTGGGTGAGAGTGAACCAAGCGTCTGAA	4259
Oy	4141	GACTGCTCAGGGCTATCCTCTCAGAGTGACATTTTAAACCACTCAGCAGAGGGATACCATG	4200
Dd	4260	GACTGCTCAGGGCTATCCTCTCAGAGTGACATTTTAAACCACTCAGCAGAGGGATACCATG	4319
Oy	4201	CAACATAACTTGATTAAGCTCCAGCAGAGAAATGGCTGACTGAAGAGCTGTGTAACAG	4260
Dd	4320	CAACATAACTTGATTAAGCTCCAGCAGAGAAATGGCTGACTGAAGAGCTGTGTAACAG	4379
Oy	4261	CATGGAGAGCCAGCCCTTCAACAGCTACCCCTCCATCAATTAAGTACTTCTGCCCCTGAG	4320
Dd	4380	CATGGAGAGCCAGCCCTTCAACAGCTACCCCTCCATCAATTAAGTACTTCTGCCCCTGAG	4439
Oy	4321	GACCTGCGAAATCCAGAAACAAGACATCAGAAAGACAGATTAACCTCCAGAAAGT	4380
Dd	4440	GACCTGCGAAATCCAGAAACAAGACATCAGAAAGACAGATTAACCTCCAGAAAGT	4499
Oy	4381	AGTGAATTCCTTATPAAGCCAGATATCCAGAAAGCCCTTCTGCTGCAAGTTTGAAGTGTCT	4440
Dd	4500	AGTGAATTCCTTATPAAGCCAGATATCCAGAAAGCCCTTCTGCTGCAAGTTTGAAGTGTCT	4559

QY	4441	GCAGATAGTCTTACCAAGTAAAAATAAGAACCCAGAGTGGAAAGGTCATCCCTTTCTAAA	4500
Db	4560	GCAGATAGTCTTACCAAGTAAAAATAAGAACCCAGAGTGGAAAGGTCATCCCTTTCTAAA	4619
QY	4501	TGCCCATCATTAGATAGATAGGTGGTACATGCAACAGTTGCTGTGGAGTCTTCACAAATAGA	4560
Db	4620	TGCCCATCATTAGATAGATAGGTGGTACATGCAACAGTTGCTGTGGAGTCTTCACAAATAGA	4679
QY	4561	AACATCCCATCTCAAGAGAGAGCTCTATTAAAGTTGGTATGGATGGAGACCAACAGCTGGAA	4620
Db	4680	AACATCCCATCTCAAGAGAGAGCTCTATTAAAGTTGGTATGGATGGAGACCAACAGCTGGAA	4739
QY	4621	GAGTCTGGGGCCACAGATTTTGACGGAAACATCTTACTTGCACAGGCAAGATCTAGAGGA	4680
Db	4740	GAGTCTGGGGCCACAGATTTTGACGGAAACATCTTACTTGCACAGGCAAGATCTAGAGGA	4799
QY	4681	ACCCCTTACCTGGGAATCTGGATACACCTCTTTCTGTATGACCTGAAATCTGATCTTCT	4740
Db	4800	ACCCCTTACCTGGGAATCTGGGAATGACGCTCTTCTGTATGACCCGTGAATCTGATCTTCT	4859
QY	4741	GAAAGCACAGCCCCAGAGTCAAGCTCGTGTGGCAACATACCATCTCAACCTCGCAATG	4800
Db	4860	GAAAGCACAGCCCCAGAGTCAAGCTCGTGTGGCAACATACCATCTCAACCTCGCAATG	4919
QY	4801	AAAGTTCCTCCCATTTGAAGTGTGACAAATCTGCCAGAGTCCAGCTGCTCATACTACT	4860
Db	4920	AAAGTTCCTCCCATTTGAAGTGTGACAAATCTGCCAGAGTCCAGCTGCTCATACTACT	4979
QY	4861	GATACCTGCTGGGTATATGCATGGAGAAAGTGTGACAGGAGAAACCCAGAATTGACA	4920
Db	4980	GATACCTGCTGGGTATATGCATGGAGAAAGTGTGACAGGAGAAACCCAGAATTGACA	5039
QY	4921	GCTTCAACAGAAAGGTCAACAAAAGATGTCCATGGGGGTGTGGGCTTACCCCGAA	4980
Db	5040	GCTTCAACAGAAAGGTCAACAAAAGATGTCCATGGGGGTGTGGGCTTACCCCGAA	5099
QY	4981	GAATTTATGCTCGGTACAAAGTTTGCCAGAAACCCACATCACTTTAACTAATCTAAT	5040
Db	5100	GAATTTATGCTCGGTACAAAGTTTGCCAGAAACCCACATCACTTTAACTAATCTAAT	5159
QY	5041	ACTGAAGAGCACTACTCATGTTGTTATGAAAACAGATGCTGAGTTGTGTGAACGGACA	5100
Db	5160	ACTGAAGAGCACTACTCATGTTGTTATGAAAACAGATGCTGAGTTGTGTGAACGGACA	5219
QY	5101	CTGAATAATTTTCTAGGAATTCGGGGAGAAATGGTAGTTACTATTTCGTGGGTGACC	5160
Db	5220	CTGAATAATTTTCTAGGAATTCGGGGAGAAATGGTAGTTACTATTTCGTGGGTGACC	5279
QY	5161	CAGTCTATTAAGAAAGAAAAATGCTGATAGATGATTTTGAAGTCAGAGAGAGATGTG	5220
Db	5280	CAGTCTATTAAGAAAGAAAAATGCTGATAGATGATTTTGAAGTCAGAGAGAGATGTG	5339
QY	5221	GTCAAATGGAGAAACCCACCAAGTCCAAAGCGAGCAAGAGATCCACAGACAGAAATGC	5280
Db	5340	GTCAAATGGAGAAACCCACCAAGTCCAAAGCGAGCAAGAGATCCACAGACAGAAATGC	5399
QY	5281	TTCAAGGGGGCTAGAAATCTGTTGCTATGAGGCCCTTCACCAACATGCCACAGATCACTG	5340
Db	5400	TTCAAGGGGGCTAGAAATCTGTTGCTATGAGGCCCTTCACCAACATGCCACAGATCACTG	5459
QY	5341	GAATGGATGGTACAGCTGTGTGGTCTCTGTGGAGGAGAGCTTTATCATCTTACACCTT	5400
Db	5460	GAATGGATGGTACAGCTGTGTGGTCTCTGTGGAGGAGAGCTTTATCATCTTACACCTT	5519
QY	5401	GGCACAGGTGTCCACCCCAATTTGTGTTGTGCACGACAGATGCTTGACAGAGACAAATGC	5460
Db	5520	GGCACAGGTGTCCACCCCAATTTGTGTTGTGCACGACAGATGCTTGACAGAGACAAATGC	5579
QY	5461	TTTCATGCAATTTGGGCGAGATGTGTGAGGACACTTGGTGACCCGAGAGTGGGTCTTGGAC	5520
Db	5580	TTTCATGCAATTTGGGCGAGATGTGTGAGGACACTTGGTGACCCGAGAGTGGGTCTTGGAC	5639

Oy		5521 AGTGTAGACACTCTACCAGGCGCAGAGCTGGACACTGCATTGAATGCCCATGCCAATCCCCAAC	5589
Dd		5640 AOTGTAGACACTCTACCAGTGCCAGAGACTGGACACTTCATGTAATCCCACAATCCCCCAC	5699
Oy	5581 AGCCCACTAC	5589       5700 AGCCAATACTAC	5708
Db	5700 AGCCAATACTAC	5708	
RESULT 5			
AAYV46463		AAV46463 standard; cDNA; 5711 BP.	
XX AC	AAV46463;		
XX XX	18-NOV-1998	(first entry)	
DE DE	Human BRCA1 omil2 polymorphism #6 CDNA.		
KW KW	BRCA1: omil2; human; breast and ovarian cancer predisposing gene; polymorphism; susceptibility; anti-oncogene; tumour suppressor; chromosome 17q; ss.		
KM KM	Homo sapiens.		
OS OS			
FH FH	Key Location/Qualifiers		
FT FT	CDS 120..5711		
FT FT	//tag= a		
FT FT	//product= "BRCA1 omil2 protein"		
FT FT	variation 4427		
FT FT	//tag= b		
FM FM	/note= "This polymorphic variation can be a T or C nucleotide"		
PN PN	US5750400-A.		
PD PD	12-MAY-1998.		
XX PF	12-FEB-1997;	97US-0798691.	
PR PR	12-FEB-1996;	96US-0598591.	
PR PR	12-FEB-1997;	97US-0798691.	
XX PA	(ONCO-) ONCORMED INC.		
PI PI	Allan AC, Alvares CP, Critz BS, Murphy PD, Olson SJ,		
PI PI	Scheelter DB, Zeng B;		
DR DR	WPt: 1998-296774/26.		
PT PT	BRCA1 oml gene coding sequences - useful for distinguishing between polymorphisms and mutation(s) in the screening for disposition to breast or ovarian cancer		
PS PS	Claim 2e; Page -: 54pp; English.		
XX CC	This sequence encodes a human BRCA1 (breast and ovarian cancer predisposing gene) omil2 gene in which a polymorphic variation occurs at nucleotide 4427. This sequence and other polymorphic variations of this sequence are useful for the identification of an individual who may or may not have an increased susceptibility to breast or ovarian cancer. The sequences used identify gene changes which are due to polymorphisms, rather than cancer causing mutations. BRCA1 is an anti-oncogene (tumour suppresser) which is involved in genetic inheritance of cancers, especially breast and ovarian cancer. It is found at human chromosome 17q which is known to be linked to cancer susceptibility, especially breast cancer. Cells containing a mutation in this gene lose the wild-type function of BRCA1 and are more susceptible to cancers. NOTE: This sequence does not appear in the specification but has been created from the wild type BRCA1 omil2 gene represented in AAV46449.		
XO XO	Sequence 5711 BP; 1956 A; 1098 C; 1274 G; 1382 T; 1 other:		

Query Match	Similarity	100.0%;	Score 5587;	DB 19;	Length 5711;
Best Local	Similarity	100.0%;	Pred. No. 0;		
Matches 5567;	Conservative	1;	Mismatches	1;	Indels 0; Gaps 0;
Qy	1	ATGATATTTATCTGCTCTTCGCGTTGAGAGATGACAAATGTCATTAAATGCTATGACGAAA	60		
Db	120	ATGATATTTATCTGCTCTTCGCGTTGAGAGATGACAAATGTCATTAAATGCTATGACGAAA	179		
Qy	61	ATCTTAGAGTGTCCATCTGTGTGGAGTTGATCAAGAACTGTCTCCAAAGATGTGAC	120		
Db	180	ATCTTAGAGTGTCCATCTGTGTGGAGTTGATCAAGAACTGTCTCCAAAGATGTGAC	239		
Qy	121	CACATATTTTGGAAATTTTGGATGCGGAAACTTTCACACGACGAAAGGGCCTTCACAG	180		
Db	240	CACATATTTTGGAAATTTTGGATGCGGAAACTTTCACACGACGAAAGGGCCTTCACAG	299		
Qy	181	TGTCCCTTATGTAAAGATGATTAACCAAAAGAGCCTCAAGAAGTACGAGATTAGT	240		
Db	300	TGTCCCTTATGTAAAGATGATTAACCAAAAGAGCCTCAAGAAGTACGAGATTAGT	359		
Qy	241	CAACTGTGTAAGAGCTATTGAAATTCATTTGTGCTTTTCAGCTTGACACAGTTTGGAG	300		
Db	360	CAACTGTGTAAGAGCTATTGAAATTCATTTGTGCTTTTCAGCTTGACACAGTTTGGAG	419		
Qy	301	TATGCAACAGCTATAATTTTGCACAAAAGGAAATTAACCTCTCCGTAACATCTAAAGAT	360		
Db	420	TATGCAACAGCTATAATTTTGCACAAAAGGAAATTAACCTCTCCGTAACATCTAAAGAT	479		
Qy	361	GAGTTTCTATCATCCAAAGTATGGGCTACAGAAACCGTGCCAAAGACCTTTCACAGAT	420		
Db	480	GAGTTTCTATCATCCAAAGTATGGGCTACAGAAACCGTGCCAAAGACCTTTCACAGAT	539		
Qy	421	GAACCCGAAATCTCTTCTTGACAGAAACCAGTCTCAGTGTCAACTCTTAACCTTGGGA	480		
Db	540	GAACCCGAAATCTCTTCTTGACAGAAACCAGTCTCAGTGTCAACTCTTAACCTTGGGA	599		
Qy	481	ACTGTGAGAACCTGTGAGAGCAAGAGGGATACAAACCTCAAAAGACCTGTCTACACTT	540		
Db	600	ACTGTGAGAACCTGTGAGAGCAAGAGGGATACAAACCTCAAAAGACCTGTCTACACTT	659		
Qy	541	GAATTGGAGTCTGATTTCTTCTGAAGATACCGTTAATAAGGCAACTTATTTGGAGTTGGGA	600		
Db	660	GAATTGGAGTCTGATTTCTTCTGAAGATACCGTTAATAAGGCAACTTATTTGGAGTTGGGA	719		
Qy	601	GATCAGAATTTGTTACAATCACCCCTCAAGGACACAGGATGAATCAGTTTGGATTCT	660		
Db	720	GATCAGAATTTGTTACAATCACCCCTCAAGGACACAGGATGAATCAGTTTGGATTCT	779		
Qy	661	GCAAAAAAGCGCTGCTTGTGAATTTTCTGAGAGGATGTAAACAATTAAGCAATCATCAAA	720		
Db	780	GCAAAAAAGCGCTGCTTGTGAATTTTCTGAGAGGATGTAAACAATTAAGCAATCATCAAA	839		
Qy	721	CCAGATTAATAATGATTTGAACACCCTGAGAAAGCGTGAAGCTGAGAGGATCTCCAGAAAG	780		
Db	840	CCAGATTAATAATGATTTGAACACCCTGAGAAAGCGTGAAGCTGAGAGGATCTCCAGAAAG	899		
Qy	781	TATCAGGAGTCTGTTTCAAACCTTGCATGTGAGAGCCATGTGGCACAAATACTCATGCC	840		
Db	900	TATCAGGAGTCTGTTTCAAACCTTGCATGTGAGAGCCATGTGGCACAAATACTCATGCC	959		
Qy	841	AGCTCATTAACAGCATGAGACAGCAGTTTATTACTCACTAAAGACAGATTAATGTAGAA	900		
Db	960	AGCTCATTAACAGCATGAGACAGCAGTTTATTACTCACTAAAGACAGATTAATGTAGAA	1019		
Qy	901	AAGGCTGAATTTCTGTAAATAAAGCAAAAGCGCTGCTTAGCAAGAGCAATCAACAGA	960		
Db	1020	AAGGCTGAATTTCTGTAAATAAAGCAAAAGCGCTGCTTAGCAAGAGCAATCAACAGA	1079		
Qy	961	TGGGCTGGAAGTAAAGAAAACATGTAATGATAGGGGAGCTCCACACAGAAAAAAGGTA	1020		
Db	1080	TGGGCTGGAAGTAAAGAAAACATGTAATGATAGGGGAGCTCCACACAGAAAAA			

1140 GATCTGATCTGATCCCTGCTGAGAGAAAAGAAATGAAATAGCAAACTCCATGCG 1199  
1081 TCGAGAAATCTGAGATACGGAAGATGTCCTGGATTAACACATAATAGCAGCATTCAG 1140  
1200 TCGAGAAATCTGAGATACGGAAGATGTCCTGGATTAACACATAATAGCAGCATTCAG 1259  
1141 AAATGATAGTGGTTTTCAGAAATGATGAATGTTAGTTCGATGACTCATGAT 1200  
1260 AAAGTTAATAGTGGTTTTCAGAAATGATGAATGTTAGTTCGATGACTCATGAT 1319  
1201 GGGAGTCTGAATCAAAATGCCAAAGTAGCTGATGATTTGACGTTCTAAATGAGTAGAT 1260  
1320 GGGAGTCTGAATCAAAATGCCAAAGTAGCTGATGATTTGACGTTCTAAATGAGTAGAT 1379  
1261 GAATATCTGTTCTTTCAGAAATGATGATGTTAGTTCGATGACTCATGAT 1320  
1380 GAATATCTGTTCTTTCAGAAATGATGATGTTAGTTCGATGACTCATGAT 1439  
1321 ATATGTAAGTGAAGAGTTCACCTCCAAATCAGTAGAGATATATTGAAGACAAATA 1380  
1440 ATATGTAAGTGAAGAGTTCACCTCCAAATCAGTAGAGATATATTGAAGACAAATA 1499  
1381 TTTTGGGAAAACCTATCGGAAGAGGCAAGCCTCCCACTTAAGCCATGTAAGTGAAT 1440  
1500 TTTTGGGAAAACCTATCGGAAGAGGCAAGCCTCCCACTTAAGCCATGTAAGTGAAT 1559  
1441 CTATATATAGAGATTTGTTACTGAGCCACAGATATATACAAAGGCTCCCTCCACAAAT 1500  
1560 CTATATATAGAGATTTGTTACTGAGCCACAGATATATACAAAGGCTCCCTCCACAAAT 1619  
1501 AAATTAAGGCTTAAGAGAGACCTACATCAGGCTTCATCCTGAGATTTTATCAAGAAA 1560  
1620 AAATTAAGGCTTAAGAGAGACCTACATCAGGCTTCATCCTGAGATTTTATCAAGAAA 1679  
1561 GCAGATTTGGCAGTTCAAAAGACTCCTGAATGATTAATCAGGAGACTAACCAACGAG 1620  
1680 GCAGATTTGGCAGTTCAAAAGACTCCTGAATGATTAATCAGGAGACTAACCAACGAG 1739  
1621 CAGATGCTCAAGTGAATGATTAATAGTGGTATGAGTAATCAAAAGGATGAT 1680  
1740 CAGATGCTCAAGTGAATGATTAATAGTGGTATGAGTAATCAAAAGGATGAT 1799  
1681 TCTATTCAGATGAGAAAATCTCTAACCAATAGAAATCAGTGAAGAAATGCTTTC 1740  
1800 TCTATTCAGATGAGAAAATCTCTAACCAATAGAAATCAGTGAAGAAATGCTTTC 1859  
1741 AAAACGAAAGCTGACCTATTAAGCAGATTAAGCAATATGCAATTCGAATTAATATC 1800  
1860 AAAACGAAAGCTGACCTATTAAGCAGATTAAGCAATATGCAATTCGAATTAATATC 1919  
1801 CACAATTCAAAGAGACCTTAAGAAATAGGCTGAGGAGAGTCTTACAGGCAATAT 1860  
1920 CACAATTCAAAGAGACCTTAAGAAATAGGCTGAGGAGAGTCTTACAGGCAATAT 1979  
1861 CATGCGCTTAAGTGAATGATTAAGCAATATGCAATTCGAATTAATATC 1920  
1980 CATGCGCTTAAGTGAATGATTAAGCAATATGCAATTCGAATTAATATC 2039  
1921 ATGATATGTTGTTCTGACAGTGAAGAGATTAAGAAAAAGTCAACAAATGCGAGTC 1980  
2040 ATGATATGTTGTTCTGACAGTGAAGAGATTAAGAAAAAGTCAACAAATGCGAGTC 2099  
1981 AGGACAGCAGAAACCTTAAGCAATCATGGAAGTAAAGAACCTGCAATGAGGCAAGAG 2040  
2100 AGGACAGCAGCAGAAACCTTAAGCAATCATGGAAGTAAAGAACCTGCAATGAGGCAAGAG 2159  
2041 AGTAACAGCAGCAATGAAGAGAGATTAAGAGACATGACAGTATATTTCCAGAGCTG 2100  
2160 AGTAACAGCAGCAATGAAGAGAGATTAAGAGACATGACAGTATATTTCCAGAGCTG 2219  
2101 AAGTTAACAATGACCTGCTTTTACTAGTTCGAATACCAAGTGAATTAAGAA 2160

2220 AAGTTAACAATGACCTGCTTTTACTAGTTCGAATTAAGTGAATTAAGAA 2279  
2161 TTTGCAATCTTACCTTCCAGAGAGAGAAAAGAGAAATAGTAAGTAAAGTGAAGTGC 2220  
2280 TTTGCAATCTTACCTTCCAGAGAGAGAAAAGAGAAATAGTAAGTAAAGTGAAGTGC 2339  
2221 TCTAATATGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2280  
2340 TCTAATATGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2339  
2281 GAAGATCTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2340  
2400 GAAGATCTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2459  
2341 GAAGATCTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2400  
2460 GAAGATCTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2519  
2401 TGTGTAGTCAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2460  
2520 TGTGTAGTCAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2579  
2461 GATATATGAATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520  
2580 GATATATGAATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2639  
2521 CGGGAACAGCAGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580  
2640 CGGGAACAGCAGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2699  
2581 TTTCAAGTTTCAAGGCGCAGATGATTTGCTGCTGTTTCAATCCAGAGAAATGCGAGAG 2640  
2700 TTTCAAGTTTCAAGGCGCAGATGATTTGCTGCTGTTTCAATCCAGAGAAATGCGAGAG 2759  
2641 GAATGTCACATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700  
2760 GAATGTCACATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2819  
2701 TTTGAATGTGAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2760  
2820 TTTGAATGTGAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2879  
2761 CAGACATTAATATCAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2820  
2880 CAGACATTAATATCAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2939  
2821 AATGCCAATGTATGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2880  
2940 AATGCCAATGTATGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2999  
2881 AAGCAACGTGAGACATTAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2940  
3000 AAGCAACGTGAGACATTAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3059  
2941 CCAACATTTTCCATCAAGTCAATTTGTTAAATCAAGTCAATTTGTTAAATCAAGTCAAG 3000  
3060 CCAACATTTTCCATCAAGTCAATTTGTTAAATCAAGTCAATTTGTTAAATCAAGTCAAG 3119  
3001 GAAACCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060  
3120 GAAACCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3179  
3061 AGTAACAGTGAAGCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3120  
3180 AGTAACAGTGAAGCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3239  
3121 TCAAGCAATTTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3180  
3240 TCAAGCAATTTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3299  
3181 ATAGGTTCCAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3240  
3300 ATAGGTTCCAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3359

QY	3241	AATGCTATGCTTAATTTAGGGGTTTTCGAACCTGAGGCTCTAAACAAGCTTCCTCGA	3300
Db	3360	AATGCTATGCTTAGATTAGGGGTTTTTGCACCTGAGGCTCTAAACAAGCTTCCTCGA	3419
QY	3301	AGTAATTGTAAAGCATCCTCGAATTAAGAAAGCAAGATTTGAAGAAGTATGTCACACTGTT	3360
Db	3420	AGTAATTGTAAAGCATCCTCGAATTAAGAAAGCAAGATTTGAAGAAGTATGTCACACTGTT	3479
QY	3361	AATACAGATTTCTCTCCATATCTGATTTACATTAATTTAGAACACCTATGGGAAGTATG	3420
Db	3480	AATACAGATTTCTCTCCATATCTGATTTACATTAATTTAGAACACCTATGGGAAGTATG	3539
QY	3421	CATACATCTCAGGTTTGTTGTCGAGACACCTGATACCTGTTAGATGATGGGAATTAAG	3480
Db	3540	CATGCAATCTCAGGTTTGTTGTCGAGACACCTGATACCTGTTAGATGATGGGAATTAAG	3539
QY	3481	GAGATATCTAGTTTGGCTGAAATATGACATTAAAGAAAGTCTGCTGTTTTAGCAAAAGC	3540
Db	3600	GAGATATCTAGTTTGGCTGAAATATGACATTAAAGAAAGTCTGCTGTTTTAGCAAAAGC	3659
QY	3541	GTCGAGAAGAGAGCTTAGCAGGAGTCTAGCCCTTTCACCCATPACATATTGGGCTCAG	3600
Db	3660	GTCGAGAAGAGAGCTTAGCAGGAGTCTAGCCCTTTCACCCATPACATATTGGGCTCAG	3719
QY	3601	GGTTACCGAAGAGGGGCCAAGAAATTTAGAGTCTCAGAAAGAACTTATCTAGTAGAGAT	3660
Db	3720	GGTTACCGAAGAGGGGCCAAGAAATTTAGAGTCTCAGAAAGAACTTATCTAGTAGAGAT	3779
QY	3661	GAAAGAGCTCCCTCTTCCACACTTGTATTGTTGGTAAAGTAAATATACCTTTCACAG	3720
Db	3780	GAAAGAGCTCCCTCTTCCACACTTGTATTGTTGGTAAAGTAAATATACCTTTCACAG	3839
QY	3721	TCTACTAGGAGCTAGCACCGTTGCTACGAGTGTGTCTAGAAACACAGAGAGAAATTTA	3780
Db	3840	TCTACTAGGAGCTAGCACCGTTGCTACGAGTGTGTCTAGAAACACAGAGAGAAATTTA	3899
QY	3781	TTATCATTTGAAGAATAGCTTAATATGACTGCAATACCAAGTAAATTTGGCAAGGCAATCT	3840
Db	3900	TTATCATTTGAAGAATAGCTTAATATGACTGCAATACCAAGTAAATTTGGCAAGGCAATCT	3959
QY	3841	CAGGAACATCACCTTAGTGAGAGAAACAAATGTTCTGCTACTGTTTTCCTTCACACTGC	3900
Db	3960	CAGGAACATCACCTTAGTGAGAGAAACAAATGTTCTGCTACTGTTTTCCTTCACACTGC	4019
QY	3901	AGTGAATTTGGAAGACTTGACTGCAATATCAACAACCCAGAGATCCCTTCTTGATTGGTTCT	3960
Db	4020	AGTGAATTTGGAAGACTTGACTGCAATATCAACAACCCAGAGATCCCTTCTTGATTGGTTCT	4079
QY	3961	TCCAAACCAAAATGAGGCAATCAGTCTGAAAGCCAGGAGATGGTCTGAGTGCACAGGAATTG	4020
Db	4080	TCCAAACCAAAATGAGGCAATCAGTCTGAAAGCCAGGAGATGGTCTGAGTGCACAGGAATTG	4139
QY	4021	GTTTTCAGATGATGAAAGAAAGGAAGGAGGCTTTGGAAGAAATAATCAAGAAAGCAAAAGC	4080
Db	4140	GTTTTCAGATGATGAAAGAAAGGAAGGAGGCTTTGGAAGAAATAATCAAGAAAGCAAAAGC	4199
QY	4081	ATGCAATTCAAACTTAGTGTGAAGGACGCAATCTGGGTGTGAGAGTGAACAACGCTCTTGAA	4140
Db	4200	ATGCAATTCAAACTTAGTGTGAAGGACGCAATCTGGGTGTGAGAGTGAACAACGCTCTTGAA	4259
QY	4141	GACGCTCAGAGGCTATCTCTCACAAGTACATTTTAACCACTCGACAGAGGGAATCCATG	4200
Db	4260	GACGCTCAGAGGCTATCTCTCACAAGTACATTTTAACCACTCGACAGAGGGAATCCATG	4319
QY	4201	CAACATAACCTGATTAAGCTCCACAGAGAAATGGCTGAACATAGAAAGCTGTGTTAGAACAG	4260
Db	4320	CAACATAACCTGATTAAGCTCCACAGAGAAATGGCTGAACATAGAAAGCTGTGTTAGAACAG	4379
QY	4261	CATGGGAGCCAGGCTTCTAACAAGTATACCTTCCATCTAATAGTACTCTTCTGCCCTTGAG	4320
Db	4380	CATGGGAGCCAGGCTTCTAACAAGTATACCTTCCATCTAATAGTACTCTTCTGCCCTTGAG	4439

OY	4321	GACCTGGCAATTC	CGAACAAGC	ACATCAG	AAAAAG	AGATTTACTT	CACAG	AAAGT	4380
Db	4440	GACCTGGCAAAATCC	GAAGCAAGC	ACATCAG	AAAAAG	AGATTTACTT	CACAG	AAAGT	4499
OY	4381	AGTGAATACCC	ATAGCCAG	AATCCAG	AAAGCCCTT	TGCGTAC	AACTTTGAG	GTCT	4448
Db	4500	AGTGAATACCC	ATATAGCC	AGATCCAG	AAAGCCCTT	TGCGTAC	AACTTTGAG	GTCT	4555
OY	4441	GCAGATAGTTCT	ACCAGTAAAA	ATAAAG	ACCAGAG	GTGAAAG	GTCACTCCCTT	CTAAA	4500
Db	4560	GCAGATAGTTCT	ACCAGTAAAA	ATAAAG	ACCAGAG	GTGAAAG	GTCACTCCCTT	CTAAA	4619
OY	4501	TGCCCATATTA	ATAGATAG	ATAGTGGTAC	ATGCAC	ATCTCTCT	GGGAGTCTT	CAGAA	4566
Db	4620	TGCCCATATTA	ATAGATAG	ATAGTGGTAC	ATGCAC	ATCTCTCT	GGGAGTCTT	CAGAA	4679
OY	4561	AACTACCCATCT	CAGAGAG	AGCTCTATT	TAAGTGTG	TATGTG	TGAGAG	CAACAGCTG	4620
Db	4680	AACTACCCATCT	CAGAGAG	AGCTCTATT	TAAGTGTG	TATGTG	TGAGAG	CAACAGCTG	4739
OY	4621	GAGTCTGGGCC	CACAGATTTG	ACGGAAC	ATCTTACTT	TGCCAAG	CAAGATCT	TACAGGA	4680
Db	4740	GAGTCTGGGCC	CACAGATTTG	ACGGAAC	ATCTTACTT	TGCCAAG	CAAGATCT	TACAGGA	4799
OY	4681	ACCCCTTACCT	GGAAATCT	GTGAATC	AGCTCTTCT	TGTATG	ACCTG	GAATCTG	4740
Db	4800	ACCCCTTACCT	GGAAATCT	GTGAATC	AGCTCTTCT	TGTATG	ACCTG	GAATCTG	4859
OY	4741	GAAGACAGAG	CCCCAG	AGTCA	AGCTCTGT	TGGCA	CAATAC	CACTCTT	4800
Db	4860	GAAGACAGAG	CCCCAG	AGTCA	AGCTCTGT	TGGCA	CAATAC	CACTCTT	4919
OY	4801	AAAGTTCCCA	ATTGAAAG	TTGCAG	AATCTGCC	CAAGTCC	AGCTGCT	CTATACT	4860
Db	4920	AAAGTTCCCA	ATTGAAAG	TTGCAG	AATCTGCC	CAAGTCC	AGCTGCT	CTATACT	4979
OY	4861	GATAC	TCTGGGTAT	ATATG	CAATG	TGAGAG	AAAGTGT	GAGCAGAG	4920
Db	4980	GATAC	TCTGGGTAT	ATATG	CAATG	TGAGAG	AAAGTGT	GAGCAGAG	5039
OY	4921	GCTTCAAC	AGAGGGTCA	CAACAA	AGATGTCT	CAATGGTGT	CGCTCA	ACCAGAA	4980
Db	5040	GCTTCAAC	AGAGGGTCA	CAACAA	AGATGTCT	CAATGGTGT	CGCTCA	ACCAGAA	5099
OY	4981	GAATTTAT	GCCTGTAC	AAAGTTTCC	ACGAAAC	ACCA	CATCACTT	TAATCTA	5040
Db	5100	GAATTTAT	GCCTGTAC	AAAGTTTCC	ACGAAAC	ACCA	CATCACTT	TAATCTA	5159
OY	5041	ACTGAAG	AGACTACT	CACTGTTG	TATG	AAACAC	ATGCTG	ATGTGTG	5100
Db	5160	ACTGAAG	AGACTACT	CACTGTTG	TATG	AAACAC	ATGCTG	ATGTGTG	5219
OY	5101	CTGA	AAATATTTCT	GTG	AATTCG	GGAGG	AAATAG	AGTATTTCT	5160
Db	5220	CTGA	AAATATTTCT	GTG	AATTCG	GGAGG	AAATAG	AGTATTTCT	5279
OY	5161	CAGTCTAT	TAAAGAA	AAAAATGCT	GTG	ATGAG	CATGATTTT	TGAAGT	5220
Db	5280	CAGTCTAT	TAAAGAA	AAAAATGCT	GTG	ATGAG	CATGATTTT	TGAAGT	5339
OY	5221	GTCA	ATGAG	AAACAC	CAAGG	GTCC	AGAG	AGATCC	5280
Db	5340	GTCA	ATGAG	AAACAC	CAAGG	GTCC	AGAG	AGATCC	5399
OY	5281	TTTC	AGGGGCT	GTAG	AAATCTG	TGCTG	ATG	GGGCTT	5340
Db	5400	TTTC	AGGGGCT	GTAG	AAATCTG	TGCTG	ATG	GGGCTT	5459
OY	5341	GAATG	AGATG	GTAC	AGCTGTG	TGTG	CTCTG	TG	5400
Db	5460	GAATG	AGATG	GTAC	AGCTGTG	TGTG	CTCTG	TG	5519
OY	5401	GGCA	AGGTG	CCAC	CCCAATTTG	TGGTGTG	ACGCC	ACAGTGC	5460

```

Db      5520 GGCACAGCTGTCCACCCCAATGTGTTGTGTCAGCAGCAATCTCCGTGAGACAGAGCAAAATGCGC 5579
Oy      5461 TTCCATCAATTTGGGCGATGTGTGAGGCACTGTGTGAGCCGAGAGTGGTGTGGAC 5520
Db      5580 TTCCATCAATTTGGGCGATGTGTGAGGCACTGTGTGAGCCGAGAGTGGTGTGGAC 5639
Oy      5521 AGTGTACACTCTACAGTGTGAGGCACTGTGTGAGCCGAGAGTGGTGTGGAC 5580
Db      5640 AGTGTACACTCTACAGTGTGAGGCACTGTGTGAGCCGAGAGTGGTGTGGAC 5699
Oy      5581 AGCCACTAC 5589
Db      5700 AGCCACTAC 5708

RESULT 6
AAV46464 ID AAV46464 standard; cDNA; 5711 BP.
AAV46464 AC
XX      18-NOV-1998 (first entry)
XX      Human BRCA1 omi2 polymorphism #7 cDNA.
XX      BRCA1: omi1: human: breast and ovarian cancer predisposing gene;
XX      polymorphism: susceptibility; anti-oncogene; tumour suppressor;
XX      chromosome 17q: ss.
XX      Homo sapiens.
XX      OS
XX      Key Location/Qualifiers
XX      CDS 120..5711
XX      FT /*tag- a
XX      FT /*product= "BRCA1 omi2 protein"
XX      FT variation 4956
XX      FT /*tag- b
XX      FT /*note= "this polymorphic variation can be an A or G
XX      nucleotide"
XX      US5750400-A.
XX      PD 12-MAY-1998.
XX      PE 12-FEB-1997; 97US-0798691.
XX      PR 12-FEB-1996; 96US-0598591.
XX      PR 12-FEB-1997; 97US-0798691.
XX      PA (ONCO-) ONCOMED INC.
XX      PI Allen AC, Alvares CP, Critz BS, Murphy PD, Olson SJ;
XX      PI Scheller DB, Zeng B;
XX      DR WPI; 1998-296774/26.
XX      PT BRCA1 omi1 gene coding sequences - useful for distinguishing between
XX      PT polymorphisms and mutation(s) in the screening for disposition to
XX      PT breast or ovarian cancer
XX      PS Claim 2e; Page -: 54pp; English.
XX      CC This sequence encodes a human BRCA1 (breast and ovarian cancer
XX      CC predisposing gene) omi2 gene in which a polymorphic variation occurs at
XX      CC nucleotide 4956. This sequence and other polymorphic variations of this
XX      CC sequence are useful for the identification of an individual who may or
XX      CC may not have an increased susceptibility to breast or ovarian cancer.
XX      CC The sequences used identify gene changes which are due to polymorphisms,
XX      CC rather than cancer causing mutations. BRCA1 is an anti-oncogene (tumour
XX      CC suppressor) which is involved in genetic inheritance of cancers,
XX      CC especially breast and ovarian cancer. It is found at human chromosome
XX      CC 17q which is known to be linked to cancer susceptibility, especially

```

```

CC breast cancer. Cells containing a mutation in this gene lose the
CC wild-type function of BRCA1 and are more susceptible to cancers.
CC NOTE: This sequence does not appear in the specification but has been
CC created from the wild type BRCA1 omi2 gene represented in AAV46449.
XX
SQ      Sequence 5711 BP; 1955 A; 1098 C; 1274 G; 1383 T; 1 other;
Query Match      100.0%; Score 5587; DB 19; Length 5711;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5587; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Oy      1 ATGGAATTTATCTGCTCTGCGGTTGAAGAATACAAAATGCTAATGCTATGTCAGAAA 60
Db      120 ATGGAATTTATCTGCTCTGCGGTTGAAGAATACAAAATGCTAATGCTATGTCAGAAA 179
Oy      61 ATCTTAGAGTGTCCCATCTGCTGTGAGATTGATCAAGAACCTGTCTCCACAAATGTGAC 120
Db      180 ATCTTAGAGTGTCCCATCTGCTGTGAGATTGATCAAGAACCTGTCTCCACAAATGTGAC 239
Oy      121 CACATATTTGCAAAATTTTGCATGCTGAAACTTCTCAACGAGAAAGAGGCTTCACAG 180
Db      240 CACATATTTGCAAAATTTTGCATGCTGAAACTTCTCAACGAGAAAGAGGCTTCACAG 299
Oy      181 TGTCTTTATGTAGAATGATATATACCAAAAGAGCCTACAGAAAGTACAGATTAGT 240
Db      300 TGTCTTTATGTAGAATGATATATACCAAAAGAGCCTACAGAAAGTACAGATTAGT 359
Oy      241 CAACCTTTGTAAGAGCATTTGAAATCATTTGCTTTTACGCTTGACACAGTTTGGAG 300
Db      360 CAACCTTTGTAAGAGCATTTGAAATCATTTGCTTTTACGCTTGACACAGTTTGGAG 419
Oy      301 TATGCAAAACAGCTATATTTTGCAGAAAGAAATATCTCTCTGAACATCTAAAGAT 360
Db      420 TATGCAAAACAGCTATATTTTGCAGAAAGAAATATCTCTCTGAACATCTAAAGAT 479
Oy      361 GAAGTTTCTATCATCAACAGATATGAGGCTACAGAAACCGTGCAGAAAGACTTTACAGAGT 420
Db      480 GAAGTTTCTATCATCAACAGATATGAGGCTACAGAAACCGTGCAGAAAGACTTTACAGAGT 539
Oy      421 GAACCCGAAATCTCTCTGTCGAGAAACACAGTCTCATGCTCTTAACCTTGA 480
Db      540 GAACCCGAAATCTCTCTGTCGAGAAACACAGTCTCATGCTCTTAACCTTGA 599
Oy      481 ACTGTGAACTCTGAGAGCAAAAGCAGCGGTACACCTCAAAAGAGCTGTCTATATT 540
Db      600 ACTGTGAACTCTGAGAGCAAAAGCAGCGGTACACCTCAAAAGAGCTGTCTATATT 659
Oy      541 GAATGGGATCTGATTTCTTGTGAAGATACCGTTAATAGGCAACTTATGCGAGTGGGA 600
Db      660 GAATGGGATCTGATTTCTTGTGAAGATACCGTTAATAGGCAACTTATGCGAGTGGGA 719
Oy      601 GATCAGCAATTTGTAACAATACCCCTCAAGAGAACCGAGTGAATTCAGTTGGATTCT 660
Db      720 GATCAGCAATTTGTAACAATACCCCTCAAGAGAACCGAGTGAATTCAGTTGGATTCT 779
Oy      661 GCAAAAAGAGCTGTGTGAAATTTTGTGAGAGGAGTATCAAAATGATGAACATCATCAA 720
Db      780 GCAAAAAGAGCTGTGTGAAATTTTGTGAGAGGAGTATCAAAATGATGAACATCATCAA 839
Oy      721 CCCAGTATATGATTTGTAACACACAGAGAGGCTCAGCTGAGAGCATCCGAGAAAG 780
Db      840 CCCAGTATATGATTTGTAACACACAGAGAGGCTCAGCTGAGAGCATCCGAGAAAG 899
Oy      781 TATCAGGAGTGTCTGTTTCAAACTGATGTGAGCCATGTGGCACAATATCTCATGCC 840
Db      900 TATCAGGAGTGTCTGTTTCAAACTGATGTGAGCCATGTGGCACAATATCTCATGCC 959
Oy      841 AGCCATTTAAGCATGAGAGAGAGAGTATTTACTCTAATAAGACAAATGAATGTGAA 900
Db      960 AGCCATTTAAGCATGAGAGAGAGAGTATTTACTCTAATAAGACAAATGAATGTGAA 1019
Oy      901 AAGGCTGAATTTCTGTAATAAAGCAACAGCTGTGCTTACCAAGAGCACAATACAGA 960

```

Db 1020 AAGGCTGAATTCGTATATAAAGCAACACCGCTGGCTTAGCCAAAGAGCCCAATTAACAGA 1079  
QY 961 TGGGCTGSAAGTAGAGAAACATGTAATGATAGGGGGGCTCCAGCACAGAAAAAAGSTA 1020  
Db 1080 TGGGCTGSAAGTAGAGAAACATGTAATGATAGGGGGGCTCCAGCACAGAAAAAAGSTA 1139  
QY 1021 GATCTGAATGCTGATCCCTGTGTGAGAGAAAGAAATGAATTAAGCAAACTGCCATGC 1080  
Db 1140 GATCTGAATGCTGATCCCTGTGTGAGAGAAAGAAATGAATTAAGCAAACTGCCATGC 1139  
QY 1081 TCAGAGAAATCTTAGAGATCTGAAGATGTTCTTGATTAACACTAAATAGCAGCTTACG 1140  
Db 1200 TCAGAGAAATCTTAGAGATCTGAAGATGTTCTTGATTAACACTAAATAGCAGCTTACG 1259  
QY 1141 AAGGTAATGAGAGGTTTTCCAGAGAGATGAACCTGTTAGTCTGTGATCAGCATGAT 1200  
Db 1260 AAGGTAATGAGAGGTTTTCCAGAGAGATGAACCTGTTAGTCTGTGATCAGCATGAT 1319  
QY 1201 GGGGAGCTGGAATCAATGCCAAAGTAGCTGATGATTTGACGTTCTTAATGAGGTAGAT 1260  
Db 1320 GGGGAGCTGGAATCAATGCCAAAGTAGCTGATGATTTGACGTTCTTAATGAGGTAGAT 1379  
QY 1261 GAATATTCCTGCTTCTCAGAGAAATAGACTTACTGGCCAGTATCTCTATGAGCTTTA 1320  
Db 1380 GAATATTCCTGCTTCTCAGAGAAATAGACTTACTGGCCAGTATCTCTATGAGCTTTA 1439  
QY 1321 ATATGTAAAGTAGAAGAGTTCACTCCCAATCAGTAGAGAGTAATTTGAAGCAAAATA 1380  
Db 1440 ATATGTAAAGTAGAAGAGTTCACTCCCAATCAGTAGAGAGTAATTTGAAGCAAAATA 1499  
QY 1381 TTTGGAAAACTTATCGGAAGAGCAAGCCCTCCCACTTAAGCCATGTAACTGAAAT 1440  
Db 1500 TTTGGAAAACTTATCGGAAGAGCAAGCCCTCCCACTTAAGCCATGTAACTGAAAT 1559  
QY 1441 CTAAATATAGAGACATTTGTTACTGAGCCAGATATATACAAGAGCCTCCCTCACAAT 1500  
Db 1560 CTAAATATAGAGACATTTGTTACTGAGCCAGATATATACAAGAGCCTCCCTCACAAT 1619  
QY 1501 AAATTAAGCGTAAAGAGAGACTTACATGAGCCCTTCTAGAGATTTTATCAAGAA 1560  
Db 1620 AAATTAAGCGTAAAGAGAGACTTACATGAGCCCTTCTAGAGATTTTATCAAGAA 1679  
QY 1561 GCAGATTTGGCAGTTCAAAAGACTCCTGAATGATAAATCAGGAGTAACCAACGGAG 1620  
Db 1680 GCAGATTTGGCAGTTCAAAAGACTCCTGAATGATAAATCAGGAGTAACCAACGGAG 1739  
QY 1621 CAGAAATGCAAGTGAATATTTACTAATAGTGTGATGAGAAATAAAGAAAGTGAT 1680  
Db 1740 CAGAAATGCAAGTGAATATTTACTAATAGTGTGATGAGAAATAAAGAAAGTGAT 1799  
QY 1681 TCTATTCAAGATGAGAAAAATCCTAAACCAATAGAAATCACTCGAAAAAAGATCTGCTTC 1740  
Db 1800 TCTATTCAAGATGAGAAAAATCCTAAACCAATAGAAATCACTCGAAAAAAGATCTGCTTC 1859  
QY 1741 AAAAGCAAGCTGAGCCTTAATAGCAGAGATATAAGCAATATGGAATCGAATTAATATC 1800  
Db 1860 AAAAGCAAGCTGAGCCTTAATAGCAGAGATATAAGCAATATGGAATCGAATTAATATC 1919  
QY 1801 CCAATTTCAAAAGCAGCTAAAGAAATAGGCTGAGAGAGAAAGCTTCTACAGGCAATAT 1860  
Db 1920 CCAATTTCAAAAGCAGCTAAAGAAATAGGCTGAGAGAGAAAGCTTCTACAGGCAATAT 1979  
QY 1861 CATGCGCTTGAATAGTACTAGTAAATCTAAAGCCACCTAATGTACTGAATTGCA 1920  
Db 1980 CATGCGCTTGAATAGTACTAGTAAATCTAAAGCCACCTAATGTACTGAATTGCA 2039  
QY 1921 ATTGATAGTTGTTAGCAGTGAAGATTAAGAAAAAAGATACAAACCAATGCCATGC 1980  
Db 2040 ATTGATAGTTGTTAGCAGTGAAGATTAAGAAAAAAGATACAAACCAATGCCATGC 2039  
QY 1981 AGGCACAGCAAAAGCTTACAACTCATGGAAGGTAAGAGAACTGCAACTGGAGCAAGAG 2040  
Db 2100 AGGCACAGCAAAAGCTTACAACTCATGGAAGGTAAGAGAACTGCAACTGGAGCAAGAG 2159  
QY 2041 AGTAAACAGCCCAATGAAACAGACAGTAAAGACATGACAGTACTTTCCAGAGCTG 2100  
Db 2160 AGTAAACAGCCCAATGAAACAGACAGTAAAGACATGACAGTACTTTCCAGAGCTG 2219  
QY 2101 AAGTTAAACAATGCACCTGGTCTTCTTACTAAGTGTCAATTAACAGTAACCTTAAGAA 2160  
Db 2220 AAGTTAAACAATGCACCTGGTCTTCTTACTAAGTGTCAATTAACAGTAACCTTAAGAA 2279  
QY 2161 TTTGTCAATCTTACCTTCCAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 2220  
Db 2280 TTTGTCAATCTTACCTTCCAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 2339  
QY 2221 TCTAATATGCTGAAGACCCCAAGATTCATGTTAAGTGAAGAAAGGTTTTCGAACCT 2280  
Db 2340 TCTAATATGCTGAAGATGACAGTAAATTCATTTGTTAGTGTGAGAAAGGTTTTCGAACCT 2399  
QY 2281 GAAAGATCTGAGAGATGACAGTAAATTCATTTGTTAGTGTGAGAAAGGTTTTCGAACCT 2340  
Db 2400 GAAAGATCTGAGAGATGACAGTAAATTCATTTGTTAGTGTGAGAAAGGTTTTCGAACCT 2459  
QY 2341 GAAAGTATCTGTTACTGGAAGTTAGCAGCTTAGGGAAGGCAAAACAGAACCAATATA 2400  
Db 2460 GAAAGTATCTGTTACTGGAAGTTAGCAGCTTAGGGAAGGCAAAACAGAACCAATATA 2519  
QY 2401 TGTGTGAGTCAAGTGTGAGATTTGAAAGCCCAAGGAGCTAATTCATGTTGTTCCAAA 2460  
Db 2520 TGTGTGAGTCAAGTGTGAGATTTGAAAGCCCAAGGAGCTAATTCATGTTGTTCCAAA 2579  
QY 2461 GATTAATGAATATGACACAGAGGCTTTAAGTATCCATTTGGAGCATGAATTAACACAGT 2520  
Db 2580 GATTAATGAATATGACACAGAGGCTTTAAGTATCCATTTGGAGCATGAATTAACACAGT 2639  
QY 2521 CGGGAACAAGCATAGAAATGGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 2580  
Db 2640 CGGGAACAAGCATAGAAATGGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 2699  
QY 2581 TTCAAGGTTTCAAGAGGCCAGTCAATTTGCTGTTTTCAAATCCAGGAATCCAGAGAG 2640  
Db 2700 TTCAAGGTTTCAAGAGGCCAGTCAATTTGCTGTTTTCAAATCCAGGAATCCAGAGAG 2759  
QY 2641 GAATGTGCAACATTTCTGCTGCCAGCTGTGGCTTTAAGCAAAAGTCCAAAGTCACT 2700  
Db 2760 GAATGTGCAACATTTCTGCTGCCAGCTGTGGCTTTAAGCAAAAGTCCAAAGTCACT 2819  
QY 2701 TTTGAATGTGAACAAAGAAAGAAAGAAATCAAGGAAAGAAAGAAAGAAAGAAAG 2760  
Db 2820 TTTGAATGTGAACAAAGAAAGAAAGAAATCAAGGAAAGAAAGAAAGAAAGAAAG 2879  
QY 2761 CAGACAGTTAATATCACTGAGGCTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2820  
Db 2880 CAGACAGTTAATATCACTGAGGCTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2939  
QY 2821 AATGCCAATGTATGTTCAAGAGAGCTTACGTTTGTCTATCATCTCAGTTTCAGAGGC 2880  
Db 2940 AATGCCAATGTATGTTCAAGAGAGAGCTTACGTTTGTCTATCATCTCAGTTTCAGAGGC 2999  
QY 2881 AAGCAAGTGAAGTCACTTCAATTAACAAATGAGCTTTCAAAACCAATATGCTATA 2940  
Db 3000 AAGCAAGTGAAGTCACTTCAATTAACAAATGAGCTTTCAAAACCAATATGCTATA 3059  
QY 2941 CCACAGCTTTTCCATCAAGTCAATTTGTTAAACATTAATGTAAGAAAAATCTGCTAGAG 3000  
Db 3060 CCACAGCTTTTCCATCAAGTCAATTTGTTAAACATTAATGTAAGAAAAATCTGCTAGAG 3119  
QY 3001 GAAACTTTGAGAGACATTCATGTCACCTGAAAGAAAGAAATGGAATGAGAACATTTCCA 3060  
Db 3120 GAAACTTTGAGAGACATTCATGTCACCTGAAAGAAAGAAATGGAATGAGAACATTTCCA 3179  
QY 3061 AGTACAGTGAAGCAATTTAGCCGTATTAACATTTAGAGAAAGTGTTTTAAGAACCCAGC 3120  
Db 3180 AGTACAGTGAAGCAATTTAGCCGTATTAACATTTAGAGAAAGTGTTTTAAGAACCCAGC 3239



QY 3121 TCAGCAATATTATTAAGTAGTTCACCTACTAATGAAGTGGCTCCAGTATTAAATGAA 3180  
Db 3240 TCAGCAATATTATTAAGTAGTTCACCTACTAATGAAGTGGCTCCAGTATTAAATGAA 3299  
QY 3181 ATAGGTTCCAGTATGAAAACATTCAGAGAACTGATGTAAGAACGAGAGGCCAAATATG 3240  
Db 3300 ATAGGTTCCAGTATGAAAACATTCAGAGAACTGATGTAAGAACGAGAGGCCAAATATG 3359  
QY 3241 AATGCTATGCTTAGATTAAGGGGTTTTCAGAACCTGAGGCTATTAACAAAGTCTCTCGA 3300  
Db 3360 AATGCTATGCTTAGATTAAGGGGTTTTCAGAACCTGAGGCTATTAACAAAGTCTCTCGA 3419  
QY 3301 AGTAATTTGATGATCCTGTAATTAAGGAAAGCAAGATATGAAGAGTACTGACAGTGT 3360  
Db 3420 AGTAATTTGATGATCCTGTAATTAAGGAAAGCAAGATATGAAGAGTACTGACAGTGT 3479  
QY 3361 AATACGATTTCTCTCCATATCTGATTTAGATTAAGTAACTTGAAGACGCTTATGGAAATAGT 3420  
Db 3480 AATACGATTTCTCTCCATATCTGATTTAGATTAAGTAACTTGAAGACGCTTATGGAAATAGT 3539  
QY 3421 CATGCTCTCAGGTTGTTCTGAGACACCTGATGACCTGTTAGATGATGATGTAATTAAG 3480  
Db 3540 CATGCTCTCAGGTTGTTCTGAGACACCTGATGACCTGTTAGATGATGATGTAATTAAG 3599  
QY 3481 GAAGATAGTATGTTTCTGTAATAATGACATTAAAGAAAGTTGCTGTTTTCAGCAAAAGC 3540  
Db 3600 GAAGATAGTATGTTTCTGTAATAATGACATTAAAGAAAGTTGCTGTTTTCAGCAAAAGC 3659  
QY 3541 GTCCAAAGAGAGGCTTACGAGAGTCCCTAGCCCTTACCCATACACATTTGGCTCAG 3600  
Db 3660 GTCCAAAGAGAGGCTTACGAGAGTCCCTAGCCCTTACCCATACACATTTGGCTCAG 3719  
QY 3601 GGTACCGAAGAGGGGCCAAGAAATAGAGTCCCTCAGAGAGAACTTATCTAGTGAAGAT 3660  
Db 3720 GGTACCGAAGAGGGGCCAAGAAATAGAGTCCCTCAGAGAGAACTTATCTAGTGAAGAT 3779  
QY 3661 GAAGAGCTTCCCTGCTTCCAAACACTTTGTAATTTGTAAGTAACAAATATACCTTCTCAG 3720  
Db 3780 GAAGAGCTTCCCTGCTTCCAAACACTTTGTAATTTGTAAGTAACAAATATACCTTCTCAG 3839  
QY 3721 TCTACTAGGCTATGACGCGTGTACCGAGTGTCTCTTAAGAACACAGAGGAATTTA 3780  
Db 3840 TCTACTAGGCTATGACGCGTGTACCGAGTGTCTCTTAAGAACACAGAGGAATTTA 3899  
QY 3781 TTATCTTGAAGAAATGTTAAATGACTGACAGTAACAGAGTAATTTGGCAAAAGCATCT 3840  
Db 3900 TTATCTTGAAGAAATGTTAAATGACTGACAGTAACAGAGTAATTTGGCAAAAGCATCT 3959  
QY 3841 CAGGAACATCACCTTAGTGAAGAAACAAATGTTCTGCTAGCTGTTTCTTCACAGTGC 3900  
Db 3960 CAGGAACATCACCTTAGTGAAGAAACAAATGTTCTGCTAGCTGTTTCTTCACAGTGC 4019  
QY 3901 AATGATTTGGAAGACTTGAAGTGAATTAACAAACACCCAGAGATCTTCTTGTGATGTTCT 3960  
Db 4020 AATGATTTGGAAGACTTGAAGTGAATTAACAAACACCCAGAGATCTTCTTGTGATGTTCT 4079  
QY 3961 TCCAAACAAATAGGAGCATAGTCTGAAAGCCAGAGGAGTGTCTGATGACAAAGAAATG 4020  
Db 4080 TCCAAACAAATAGGAGCATAGTCTGAAAGCCAGAGGAGTGTCTGATGACAAAGAAATG 4139  
QY 4021 GTTTCAGATGATGAAGAAAGAGAAAGGCTTGAAGAAATATCAAGAGAGCAAAAC 4080  
Db 4140 GTTTCAGATGATGAAGAAAGAGAAAGGCTTGAAGAAATATCAAGAGAGCAAAAC 4199  
QY 4081 AATGATTTCAACTTAGTGAAGAGATGCTGGTGTGAAGTGAAGCAAGGCTCTGAA 4140  
Db 4200 AATGATTTCAACTTAGTGAAGAGATGCTGGTGTGAAGTGAAGCAAGGCTCTGAA 4259  
QY 4141 GACTGCTCAGGGCTATCTCTCAGAGTGAATTTTAACACTCAGAGAGAGGATACCATG 4200  
Db 4260 GACTGCTCAGGGCTATCTCTCAGAGTGAATTTTAACACTCAGAGAGAGGATACCATG 4319  
QY 4201 CAACATTAACCTGATTAAGCTCCAGAGAAATGGCTGAAGTGTGTTAGAACAG 4260

Db 4320 CAACATTAACCTGATTAAGCTCCAGAGAAATGGCTGAAGTGTGTTAGAACAG 4379  
QY 4261 CATGGAGGCCAGCTTCTTAACACTACCTTCCATCATTAAGTGAACCTTGGCCCTTGAG 4320  
Db 4380 CATGGAGGCCAGCTTCTTAACACTACCTTCCATCATTAAGTGAACCTTGGCCCTTGAG 4439  
QY 4321 GACCTGGAAATCCAGAACCAAGACATCCAGAAAAAGCAGTATTAACTTCCAGAAAAAGT 4380  
Db 4440 GACCTGGAAATCCAGAACCAAGACATCCAGAAAAAGCAGTATTAACTTCCAGAAAAAGT 4499  
QY 4381 AGTGAATTAACCTTAAGCCAGAAATCCAGAAAGCCTTCTGCTGACAAAGTTGAGGTCT 4440  
Db 4500 AGTGAATTAACCTTAAGCCAGAAATCCAGAAAGCCTTCTGCTGACAAAGTTGAGGTCT 4559  
QY 4441 GCAGATAGTTCTACCGTAAATAATAAGAACCCAGAGAGTGAAGAGTATCCCTTCTAA 4500  
Db 4560 GCAGATAGTTCTACCGTAAATAATAAGAACCCAGAGAGTGAAGAGTATCCCTTCTAA 4619  
QY 4501 TGCCCATCATTAAGTATAGTGTGTACATGCAAGTGTGCTGGAGTCTTCAGAAATGA 4560  
Db 4620 TGCCCATCATTAAGTATAGTGTGTACATGCAAGTGTGCTGGAGTCTTCAGAAATGA 4679  
QY 4561 AACTACCCATCTCAAGAGAGCTCATTAAGGTTGTTGATGTGAGAGACAAAGCTGGAA 4620  
Db 4680 AACTACCCATCTCAAGAGAGCTCATTAAGGTTGTTGATGTGAGAGACAAAGCTGGAA 4739  
QY 4621 GAGTGGGGCCACAGATTTGAGGAAACATCTTACTTGGCAAGGCAAGTCTAGAGGA 4680  
Db 4740 GAGTGGGGCCACAGATTTGAGGAAACATCTTACTTGGCAAGGCAAGTCTAGAGGA 4799  
QY 4681 ACCCTTACCTGGAATCTGGAATCAGCCCTTCTCTGATGACCCCTGAATCTGATCTCT 4740  
Db 4800 ACCCTTACCTGGAATCTGGAATCAGCCCTTCTCTGATGACCCCTGAATCTGATCTCTCT 4859  
QY 4741 GAAGACAGAGCCCCAGAGTCACTGCTGTTGGCAACATACCATCTTCTGCAATG 4800  
Db 4860 GAAGACAGAGCCCCAGAGTCACTGCTGTTGGCAACATACCATCTTCTGCAATG 4919  
QY 4801 AAGGTTCCCAATTTGAAAGTTGAGAAATCGCCACAGTCCAGCTGCTCATCTACT 4860  
Db 4920 AAGGTTCCCAATTTGAAAGTTGAGAAATCGCCACAGTCCAGCTGCTCATCTACT 4979  
QY 4861 GATACGCTGGGTATATGCAATGCAAGAAAGTGTGAGCAGAGAGAACCCAGAAATGACA 4920  
Db 4980 GATACGCTGGGTATATGCAATGCAAGAAAGTGTGAGCAGAGAGAACCCAGAAATGACA 5039  
QY 4921 GCTTCAACAGAAAGGCTCAACAAAGATGTCATGCTGCTGCTGACCCAGAA 4980  
Db 5040 GCTTCAACAGAAAGGCTCAACAAAGATGTCATGCTGCTGCTGACCCAGAA 5099  
QY 4981 GAATTTATGCTGCTGATCAAGTGTGCAAGAAACACACATCATCTTAATCTAAT 5040  
Db 5100 GAATTTATGCTGCTGATCAAGTGTGCAAGAAACACACATCATCTTAATCTAAT 5159  
QY 5041 ACTGAAGAGACTACTATCTTGTATGAAGAAACGATGCTGAGTTGTGTGAAGCAGACA 5100  
Db 5160 ACTGAAGAGACTACTATCTTGTATGAAGAAACGATGCTGAGTTGTGTGTGAAGCAGACA 5219  
QY 5101 CTGAATATTTCTAGGAATTCGGGAGGAAATGGGTAGTTAGTATTTCTGGTGACC 5160  
Db 5220 CTGAATATTTCTAGGAATTCGGGAGGAAATGGGTAGTTAGTATTTCTGGTGACC 5279  
QY 5161 CAGTCTATTAAGAAAGAAATGCGATGAGCAATGATTTGAAGTCTGAGAGAGATG 5220  
Db 5280 CAGTCTATTAAGAAAGAAATGCGATGAGCAATGATTTGAAGTCTGAGAGAGATG 5339  
QY 5221 GTCAATGAAGAAACCAAGGTCCTCAAGAGCAGAGCAAGAAATCCAGAGACAGAAATGATC 5280  
Db 5340 GTCAATGAAGAAACCAAGGTCCTCAAGAGCAGAGCAAGAAATCCAGAGACAGAAATGATC 5399  
QY 5281 TTCAGGGGCTAGAAATCTGTTGCTATGGGCCCTTACCAACATGCCCACAGATCACTG 5340





QY	841	NGCCTCATTCACAGCATGAGAACAGCAGTTTATTACTCACTAAAGACAGATGAATGTAGAA	900
Db	960	AGCTCATTTACAGCAGTAGAAGAACAGCAGTTTATTACTCACTAAAGACAGATGAATGTAGAA	1019
QY	901	AAGCGTGAATTCCTGTAAATAAAGCAAAACAGCCTGGCTTAGCAAGAGGCAACATTAACGA	960
Db	1020	AAGCGTGAATTCCTGTAAATAAAGCAAAACAGCCTGGCTTAGCAAGAGGCAACATTAACGA	1079
QY	961	TGGCGTGGAGTAAGAGAAACATGTATATGATATAGCGGAGCTCCAGACAGAAAAAAGGTA	1020
Db	1080	TGGCGTGGAGTAAGAGAAACATGTATATGATATAGCGGAGCTCCAGACAGAAAAAAGGTA	1139
QY	1021	GATCGATGATGGTGATCCCCGTGGTAGAGAAAAAGAAATGAATTAACACAAACTGGCATTGC	1089
Db	1140	GATCGATGATGGTGATCCCCGTGGTAGAGAAAAAGAAATGAATTAACACAAACTGGCATTGC	1199
QY	1081	TCAGAGAAATCCTAGAGATACGTAAAGATGTCTCTGGATTAACACTAAATAGCAGATTCCAG	1140
Db	1200	TCAGAGAAATCCTAGAGATACGTAAAGATGTCTCTGGATTAACACTAAATAGCAGATTCCAG	1259
QY	1141	AAAGTAATATGAGTGGTTTCCAGAAAGATATAAGTATAGGTTCCGATGACATCAATGAT	1200
Db	1260	AAAGTAATATGAGTGGTTTCCAGAAAGATATAAGTATAGGTTCCGATGACATCAATGAT	1319
QY	1201	GGGAGTCTGATCAAAATGCCAAAGTAGCTGATGATTGAGAGCTTCTAAATAGAGGTAGAT	1260
Db	1320	GGGAGTCTGATCAAAATGCCAAAGTAGCTGATGATTGAGAGCTTCTAAATAGAGGTAGAT	1379
QY	1261	GAATATTTGTGGTTCCTCAGAGAAATATAGACTTACTGGCCAGTGATCCTCATGAGGCTTTA	1320
Db	1380	GAATATTTGTGGTTCCTCAGAGAAATATAGACTTACTGGCCAGTGATCCTCATGAGGCTTTA	1439
QY	1321	ATATGTAAAAATGAAAGAGTCACTCCCAATCAGTAGAGAGTAATATATGAAGACAAAATA	1380
Db	1440	ATATGTAAAAATGAAAGAGTCACTCCCAATCAGTAGAGAGTAATATATGAAGACAAAATA	1499
QY	1381	TTTGGGAAAAACCTATCGGAGAGGCAAGCCCTCCCACTTAAGCCATGTAACTGAAAAAT	1440
Db	1500	TTTGGGAAAAACCTATCGGAGAGGCAAGCCCTCCCACTTAAGCCATGTAACTGAAAAAT	1559
QY	1441	CTAATTATATAGAGGATTTGTATCTAGAGCCACAGATAATACAAAGAGCCGCCCTCACAAT	1500
Db	1560	CTAATTATATAGAGGATTTGTATCTAGAGCCACAGATAATACAAAGAGCCGCCCTCACAAT	1619
QY	1501	AAATTTAAAGCCTAAAGAGAGACCTACATCAGAGCCCTTCATCCTGTAGAGATTTTATCAAGAA	1560
Db	1620	AAATTTAAAGCCTAAAGAGAGACCTACATCAGAGCCCTTCATCCTGTAGAGATTTTATCAAGAA	1679
QY	1561	GCAATATTTGGCAGTTCAAAAAGACTTCCTGAATATGATAAATCAAGGAACTTAACCAAAAGAG	1620
Db	1680	GCAATATTTGGCAGTTCAAAAAGACTTCCTGAATATGATAAATCAAGGAACTTAACCAAAAGAG	1739
QY	1621	CAGATATGTCAGATGAATATTTACTTAATATGATGGTCATGAGATTAACCAAAAGGTGAT	1680
Db	1740	CAGATATGTCAGATGAATATTTACTTAATATGATGGTCATGAGATTAACCAAAAGGTGAT	1799
QY	1681	TCATTTTACGATGAGAAAAATCCTTAACCAATAGAAATCAGCTCGAAAAAGAAATTCGCTTTC	1740
Db	1800	TCATTTTACGATGAGAAAAATCCTTAACCAATAGAAATCAGCTCGAAAAAGAAATTCGCTTTC	1859
QY	1741	AAAAAGAAAGCTGAACCTTAAGAGAGCGATTAAGCAATATGGAACCTGGAATTTAAATATC	1800
Db	1860	AAAAAGAAAGCTGAACCTTAAGAGAGCGATTAAGCAATATGGAACCTGGAATTTAAATATC	1919
QY	1801	CACAATTCAAAAGACCTTAAGAAATATAGCTGAGAGAGAGTCTTCTACACAGCATATT	1860
Db	1920	CACAATTCAAAAGACCTTAAGAAATATAGCTGAGAGAGAGTCTTCTACACAGCATATT	1979
QY	1861	CATCGCGCTGGAACATGATGTCAGTAAGAAATCTAAGCCACCTAATTGTACTGAATTCGAA	1920
Db	1980	CATCGCGCTGGAACATGATGTCAGTAAGAAATCTAAGCCACCTAATTGTACTGAATTCGAA	2039

QY	1921	ATTGATTA	CTTGTCTT	CAGACGTG	AAGCAAGT	TAAGA	AAAAA	AAAGTCA	ACCAAT	GGCAGT	1980											
Db	2040	ATTGTA	TA	CTTGTCTT	CAGACGTG	AAGCAAGT	TAAGA	AAAAA	AAAGTCA	ACCAAT	GGCAGT	2099										
QY	1981	AGCCAC	CACAGAA	AACTCT	CAAC	CTCAT	GGAG	AGTAA	GAAC	CTGCA	AGCCAA	GAAG	2040									
Db	2100	AGCCAC	CACAGAA	AACTCT	CAAC	CTCAT	GGAG	AGTAA	GAAC	CTGCA	AGCCAA	GAAG	2159									
QY	2041	AGTACA	AGCCAA	TGAA	TGACAG	ACAAGT	TAAGA	ACAT	GCAGT	GTACTT	CCAG	AGCTG	2100									
Db	2160	AGTACA	AGCCAA	TGAA	TGACAG	ACAAGT	TAAGA	ACAT	GCAGT	GTACTT	CCAG	AGCTG	2219									
QY	2101	AAGTTA	ACA	AAAT	TGCAC	CTG	CTTCTT	TCTTA	CTA	AGTGT	CTCA	AT	ACCAGT	GAAC	CTTAA	GA	2160					
Db	2220	AAGTTA	ACA	AAAT	TGCAC	CTG	CTTCTT	TCTTA	CTA	AGTGT	CTCA	AT	ACCAGT	GAAC	CTTAA	GA	2279					
QY	2161	TTTGCA	TCTAG	CTT	AGC	CTTCC	CAAG	GAAG	AAAA	AAAG	AAAG	AAAG	AAAG	AAAG	AAAG	AAAG	2220					
Db	2280	TTTGCA	TCTAG	CTT	AGC	CTTCC	CAAG	GAAG	AAAA	AAAG	AAAG	AAAG	AAAG	AAAG	AAAG	AAAG	2339					
QY	2221	TCTA	TAA	TGCT	GAA	AGAC	CCCA	AAAG	ATCT	CAT	GT	TAA	GTG	GA	AAAG	GGTTT	TC	CA	AACT	2280		
Db	2340	TCTA	TAA	TGCT	GAA	AGAC	CCCA	AAAG	ATCT	CAT	GT	TAA	GTG	GA	AAAG	GGTTT	TC	CA	AACT	2399		
QY	2281	GAA	GAT	CTG	TAG	AG	AGT	AGC	ATG	TTC	ATT	CAT	TG	TG	TG	TG	TG	TG	TG	TG	2340	
Db	2400	GAA	GAT	CTG	TAG	AG	AGT	AGC	ATG	TTC	ATT	CAT	TG	TG	TG	TG	TG	TG	TG	TG	2459	
QY	2341	GAA	GAT	CTG	TAG	AG	AGT	AGC	ATG	TTC	ATT	CAT	TG	TG	TG	TG	TG	TG	TG	TG	2400	
Db	2460	GAA	GAT	CTG	TAG	AG	AGT	AGC	ATG	TTC	ATT	CAT	TG	TG	TG	TG	TG	TG	TG	TG	2519	
QY	2401	TGT	G	TG	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A	2460
Db	2520	TGT	G	TG	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A	2579
QY	2461	GAT	AT	A	T	A	T	A	T	A	T	A	T	A	T	A	T	A	T	A	T	2520
Db	2580	GAT	AT	A	T	A	T	A	T	A	T	A	T	A	T	A	T	A	T	A	T	2639
QY	2521	CGG	GA	AA	CA	AC	AT	GA	AA	TG	GA	AA	GA	AA	TG	GA	AA	GA	AA	TG	GA	2580
Db	2640	CGG	GA	AA	CA	AC	AT	GA	AA	TG	GA	AA	GA	AA	TG	GA	AA	GA	AA	TG	GA	2699
QY	2581	TTT	CA	G	CTT	CA	A	G	C	G	C	A	G	C	A	G	C	A	G	C	A	2640
Db	2700	TTT	CA	G	CTT	CA	A	G	C	G	C	A	G	C	A	G	C	A	G	C	A	2759
QY	2641	GA	T	T	G	C	A	C	A	T	T	C	T	C	T	C	T	C	T	C	T	2700
Db	2760	GA	T	T	G	C	A	C	A	T	T	C	T	C	T	C	T	C	T	C	T	2819
QY	2701	TTT	GA	A	T	G	A	T	G	A	A	G	A	A	A	A	A	A	A	A	A	2760
Db	2820	TTT	GA	A	T	G	A	T	G	A	A	G	A	A	A	A	A	A	A	A	A	2879
QY	2761	CAG	AC	AG	T	A	T	A	T	A	T	A	T	A	T	A	T	A	T	A	T	2820
Db	2880	CAG	AC	AG	T	A	T	A	T	A	T	A	T	A	T	A	T	A	T	A	T	2939
QY	2821	AAT	G	C	CA	A	A	T	G	A	T	G	A	T	G	A	T	G	A	T	G	2880

```
Db 3120 GAAACATTTGAGGAGACATTCATCATGTCAACCTGAAGAAATGGAAATGAGAACATCTCA 3179
Qy 3061 AGTACAGTGGACCAATTTAGCCGTAATACATAGAGAAATGTTTTTAAGAAGCCAGC 3120
Db 3180 AGTACAGTGGACCAATTTAGCCGTAATACATAGAGAAATGTTTTTAAGAAGCCAGC 3239
Qy 3121 TCAGCAATATTTAATGAAGTAGTTCAGTACTAATGAAGTGGGCTCCAGTATTAATGA 3180
Db 3240 TCAGCAATATTTAATGAAGTAGTTCAGTACTAATGAAGTGGGCTCCAGTATTAATGA 3299
Qy 3181 ATAGGTTCCAGTGTATGAAACATTTCAAGCAAGTGTAGTGAAGAGAGGCAAAATG 3240
Db 3300 ATAGGTTCCAGTGTATGAAACATTTCAAGCAAGTGTAGTGAAGAGAGGCAAAATG 3359
Qy 3241 AATGCTATGCTTATGATTTAGGGGTTTTGCAACCTGAGTCTATTAACAAATCTTCTCGA 3300
Db 3360 AATGCTATGCTTATGATTTAGGGGTTTTGCAACCTGAGTCTATTAACAAATCTTCTCGA 3419
Qy 3301 AGTAATTTGAAGCATCTGTAATAAATAAAGCAAGATATGAAGAGTTCAGACTGT 3360
Db 3420 AGTAATTTGAAGCATCTGTAATAAATAAAGCAAGATATGAAGAGTTCAGACTGT 3479
Qy 3361 AATPACGATTTCTCTCCATATCTGATTTAGATTAAGACAGCCTATGGAAGTAGT 3420
Db 3480 AATPACGATTTCTCTCCATATCTGATTTAGATTAAGACAGCCTATGGAAGTAGT 3539
Qy 3421 CATGCATCTCAGGTTTTGTTCTGAGACACCTGATGACCTGTATGATGATGATGAATTAAG 3480
Db 3540 CATGCATCTCAGGTTTTGTTCTGAGACACCTGATGACCTGTATGATGATGATGAATTAAG 3599
Qy 3481 GAAGATACTAGTTTGTGTAATAATGACATTAAGAAAGTTCTGCTGTTTTTACAAAGC 3540
Db 3600 GAAGATACTAGTTTGTGTAATAATGACATTAAGAAAGTTCTGCTGTTTTTACAAAGC 3659
Qy 3541 GTCCAGAAAGAGAGCTTATGACAGAGTCTAGGCCCTTACCCATACACTTTGGCTCAG 3600
Db 3660 GTCCAGAAAGAGAGCTTATGACAGAGTCTAGGCCCTTACCCATACACTTTGGCTCAG 3719
Qy 3601 GGTACCCGAGAGGGGCCAAGAAATAGAGTCTCTGAGAGAGAACTTATCTAGTGAAGAT 3660
Db 3720 GGTACCCGAGAGGGGCCAAGAAATAGAGTCTCTGAGAGAGAACTTATCTAGTGAAGAT 3779
Qy 3721 TCTACTAGGCAATGACCGTGTCTACCGAGTGTCTGTCTTAAGAACACAGAGGAAATTTA 3780
Db 3780 TCTACTAGGCAATGACCGTGTCTACCGAGTGTCTGTCTTAAGAACACAGAGGAAATTTA 3839
Qy 3781 TTATCATTGAAGAAATAGCTTAATATGACTCAGTAACAGAGTAATTTGCAAAAGCATCT 3840
Db 3900 TTATCATTGAAGAAATAGCTTAATATGACTCAGTAACAGAGTAATTTGCAAAAGCATCT 3959
Qy 3841 CAGGAACATCACCTTAGTGAGGAAACAAATGTCTGCTAGCTTGTCTTCCACAGTGC 3900
Db 3960 CAGGAACATCACCTTAGTGAGGAAACAAATGTCTGCTAGCTTGTCTTCCACAGTGC 4019
Qy 3901 AGTGAATTTGGAAGACTTGTGCTGCAAAATACAAACAGGAGATCTTCTGATTTGGTTCT 3960
Db 4020 AGTGAATTTGGAAGACTTGTGCTGCAAAATACAAACAGGAGATCTTCTGATTTGGTTCT 4079
Qy 3961 TCCAAAGCAATGAGGATCAGTCTGAAAGCCAGAGGATTTGCTGAGTGAAGGAATTTG 4020
Db 4080 TCCAAAGCAATGAGGATCAGTCTGAAAGCCAGAGGATTTGCTGAGTGAAGGAATTTG 4139
Qy 4021 GTTTCAGATGATGAAGAAAGAGAAACGGCTTTGGAAGAAATATATCAGAGAGCAAGC 4080
Db 4140 GTTTCAGATGATGAAGAAAGAGAAACGGCTTTGGAAGAAATATATCAGAGAGCAAGC 4199
Qy 4081 ATGGATTCAAACTTATGAGTGAAGCAGCATCTGGGTGTGAGAGTGAACAAGCCTCTCGAA 4140
Db 4200 ATGGATTCAAACTTATGAGTGAAGCAGCATCTGGGTGTGAGAGTGAACAAGCCTCTCGAA 4259
Qy 4141 GACTGCTCAGGGGCTATCTCTCAGAGTGAACATTTTAACCACTCAGCAGAGGATTCATG 4200
Db 4260 GACTGCTCAGGGGCTATCTCTCAGAGTGAACATTTTAACCACTCAGCAGAGGATTCATG 4319
Qy 4201 CAACATTAACCTGATTAAGCTCCAGCAGAAATGGCTGAACATGAGAGCTGTGTTAGAACAG 4260
Db 4320 CAACATTAACCTGATTAAGCTCCAGCAGAAATGGCTGAACATGAGAGCTGTGTTAGAACAG 4379
Qy 4261 CATGGAGCCAGGCTTCTTACAGCTACCTTCATCATATAGTACTCTTGTGCCCTTGAG 4320
Db 4380 CATGGAGCCAGGCTTCTTACAGCTACCTTCATCATATAGTACTCTTGTGCCCTTGAG 4439
Qy 4321 GACCTGCGAAATCCAGAACAGCATGAGAAAGAGATTTTAACTTCACAGAAAGT 4380
Db 4440 GACCTGCGAAATCCAGAACAGCATGAGAAAGAGATTTTAACTTCACAGAAAGT 4499
Qy 4381 AGTAATACCTTATTAAGCCAGATCCAGAGGCTTCTGCTGACAGATTTGAGGTCTCT 4440
Db 4500 AGTAATACCTTATTAAGCCAGATCCAGAGGCTTCTGCTGACAGATTTGAGGTCTCT 4559
Qy 4441 GCAGATAGTTCTTACAGCTAATAATAAGAACAGAGAGTGAAGAGTCAATCCCTTTAA 4500
Db 4560 GCAGATAGTTCTTACAGCTAATAATAAGAACAGAGAGTGAAGAGTCAATCCCTTTAA 4619
Qy 4501 TGCCCATCAATATGATATAGTGTGATCATGACAGTGTGCTGGAGTCTCAGATTAAGA 4560
Db 4620 TGCCCATCAATATGATATAGTGTGATCATGACAGTGTGCTGGAGTCTCAGATTAAGA 4679
Qy 4561 AACTACCATCTTCAAGAGAGGCTCATTAAGTGTGATGATGAGAGACCAACAGCTGAA 4620
Db 4680 AACTACCATCTTCAAGAGAGGCTCATTAAGTGTGATGATGAGAGACCAACAGCTGAA 4739
Qy 4621 GACTGTGGGCGCACACATTTGACGGAAACATCTTACTTGTCCAGAGCAGATCTAGAGGA 4680
Db 4740 GACTGTGGGCGCACACATTTGACGGAAACATCTTACTTGTCCAGAGCAGATCTAGAGGA 4799
Qy 4681 ACCCTTACCTGGAATCGAATCAACGCTCTCTCTGATGAGCCGTAATCTGATCTCTCT 4740
Db 4800 ACCCTTACCTGGAATCGAATCAACGCTCTCTCTGATGAGCCGTAATCTGATCTCTCTCT 4859
Qy 4741 GAAGACAGAGCCCCAGAGTCACTGCTGTGGCAACATACATCTTCAACCTCTGCAATG 4800
Db 4860 GAAGACAGAGCCCCAGAGTCACTGCTGTGGCAACATACATCTTCAACCTCTGCAATG 4919
Qy 4801 AAGGTTCCCAANTGAAGTTGACAGATCTGCCAGAGTCCAGCTGCTCATACTACT 4860
Db 4920 AAGGTTCCCAANTGAAGTTGACAGATCTGCCAGAGTCCAGCTGCTCATACTACT 4979
Qy 4861 GATACGTGTGGGTATTAATGCAATGAGAAAGTGTGACAGAGGAGAAACCGAAATTTGACA 4920
Db 4980 GATACGTGTGGGTATTAATGCAATGAGAAAGTGTGACAGAGGAGAAACCGAAATTTGACA 5039
Qy 4921 GCTTCAACGAAAGGGTCAACAAAGAAATGTCATGTGTGTCTGGCTGACCCAGAA 4980
Db 5040 GCTTCAACGAAAGGGTCAACAAAGAAATGTCATGTGTGTCTGGCTGACCCAGAA 5099
Qy 4981 GAATTTATCTGCTGTAGCAAGTTTCCAGAAACCCACATCACTTTAACTAATCTAAT 5040
Db 5100 GAATTTATCTGCTGTAGCAAGTTTCCAGAAACCCACATCACTTTAACTAATCTAAT 5159
Qy 5041 ACTGAAGACATCTACATGTTGTATGAAACAGATGCTGAGTTGTGTGTAAGCGGACA 5100
Db 5160 ACTGAAGACATCTACATGTTGTATGAAACAGATGCTGAGTTGTGTGTAAGCGGACA 5219
Qy 5101 CTGAATATTTTCTAGGAATTTGCGGAGAAATGGAGTATTAATCTATTTCTGGGTGACC 5160
Db 5220 CTGAATATTTTCTAGGAATTTGCGGAGAAATGGAGTATTAATCTATTTCTGGGTGACC 5279
Qy 5161 CAGTCTATTTAAAGAAAGAAATGCTGAATGAGCATGATTTTGAAGTCAAGAGGAGATGTG 5220
Db 5280 CAGTCTATTTAAAGAAAGAAATGCTGAATGAGCATGATTTTGAAGTCAAGAGGAGATGTG 5339
```

OY		5221	GTCATGTGAAGAAACACCACCAAGTCTCAAAGAAGACAAGTAATCCACGACGACAAAAGATC	5280				
Dd		5340	GTCATGTGAAGAAACACCACCAAGTCTCAAAGAAGACAAGTAATCCACGACGACAAAAGATC	5399				
OY		5281	TTCAGGGGGCGCTAGAATACTGTGGCTAATGGGCCCTTCACCAACATGCCCACAGATCAACTG	5340				
OY		5341	GAATGGAAATTGTACGCCTGTGGTGCTTTCTGTGGTGAAGAGCTTTTATCATTTACCCCTT	5400				
Dd		5460	GAAATGGAATGTACAGCTGTGTGGTCTTGTGTGGAAGGAGCTTTTCAATTCACCCCTT	5519				
OY		5401	GGCAGAGTGTCCACCCCAGTAATGTGGTTGTGAGCAGATAGCTGTGAGACAGAGACAATGGC	5466				
Dd		5520	GGCACAGSTGTCCACCCCAATGTGTGTTGTGACAGCCAAGTAGTCTGGACAGAGACAATBTGC	5579				
OY		5461	TTCCATGCAATTGGGCGAGATGTGTGAGCGACCCCTGTGTGACCCGAGATGGSGTGTGGAC	5520				
Dd		5580	TTCCATGCAATTGGGCGAGATGTGTGAGCGACCCCTGTGTGACCCGAGATGGSGTGTGGAC	5639				
OY		5521	AGTGTAGACCTCTACCCAGTGTGCAGAGGTGGACACCTCTACCTGATACCCCAATCCCCCAG	5588				
Dd		5640	AGTGTAGACCTCTACCCAGTGTGCAGAGGTGGACACCTCTACCTGATACCCCAATCCCCCAG	5699				
OY		5581	AGCCCACTAC 5589 					
Dd		5700	AGCCCACTAC 5708					
RESULT 8								
ID	AAV46459							
XX	AAV46459 standard; cDNA; 5711 BP.							
AC	AAV46459;							
DT	18-NOV-1998	(first entry)						
DE	Human BRCA1 omi2 polymorphism #2 cDNA.							
KW	BRCA1: omi2: human: breast and ovarian cancer predisposing gene:							
KM	polymorphism, susceptibility; anti-oncogene; tumour suppressor;							
XX	chromosome 17q; ss.							
OS	Homo sapiens.							
FH	Key	Location/Qualifiers						
FT	CDS	120..5711						
FT	/tag= a							
FT	/product= "BRCA1 omi2 protein"							
FT	variation	2430						
FT	/tag= b							
FT	/note= "This polymorphic variation can be a T or C nucleotide"							
PX	US5750400-A.							
PN								
PD	12-MAY-1998.							
PE	12-FEB-1997;	97US-0798691.						
PR	12-FEB-1996;	96US-0598591.						
PR	12-FEB-1997;	97US-0798691.						
PA	(ONCO-) ONCOMED INC.							
PI	Allien AC, Alvares CP,	Critz BS, Murphy PD, Olson SJ;						
PI	Schelter DB, Zeng B;							
DR	WPI: 1998-296774/26.							
BRCA1 omi gene coding sequences - useful for distinguishing between polymorphisms and mutation(s) in the screening for disposition to								

PT breast or ovarian cancer  
XX  
PS Claim 2e; Page -: 54pp; English.  
XX  
CC This sequence encodes a human BRCA1 (breast and ovarian cancer  
CC predisposing gene) om12 gene in which a polymorphic variation occurs at  
CC nucleotide 2430. This sequence and other polymorphic variations of this  
CC sequence are useful for the identification of an individual who may or  
CC may not have an increased susceptibility to breast or ovarian cancer.  
CC The sequences used identify gene changes which are due to polymorphisms,  
CC rather than cancer causing mutations. BRCA1 is an anti-oncogene (tumour  
CC suppressor) which is involved in genetic inheritance of cancers,  
CC especially breast and ovarian cancer. It is found at human chromosome  
CC 17q which is known to be linked to cancer susceptibility, especially  
CC breast cancer. Cells containing a mutation in this gene lose the  
CC wild-type function of BRCA1 and are more susceptible to cancers.  
CC NOTE: This sequence does not appear in the specification but has been  
CC created from the wild type BRCA1 om12 gene represented in AA046449.  
XX  
SQ Sequence 5711 BP; 1956 A; 1098 C; 1274 G; 1382 T; 1 other;

Query Match	100.0%;	Score 5587;	DB 19;	Length 5711;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 5587; Conservative	1;	Mismatches	1;	Indels 0; Gaps 0;

OY	1	TTGGATTTATCTGCTTCCTTGGCTTGAAGAAAGTACAAAAATGTAATTAATGCTATGACGAAA	60
Db	120	ATGATATTATCTGCTCTTCCGCTTGAGAAAGTACAAAATGTCATTATATGCTATGACGAAA	179
OY	61	ATCTTAGAGTGTCCCATCTGCTGAGTGTATCAAGGAAGCACTGTCCCAAAAGTGTGAC	120
Db	180	ATCTTAGAGTGTCCCATCTGTCTGTGAGTGTATCAAGGAAGCACTGTCCCAAAAGTGTGAC	239
OY	121	CACATATTTTGC AAAATTTTGCATGCTGAAACTTCTCACACGAGAAGAGGCGCTTCACAG	180
Db	240	CACATATTTTGC AAAATTTTGCATGCTGAAACTTCTCACACGAGAAGAGGCGCTTCACAG	299
OY	181	TGTCCTTATGTAAAGATATTAACCAAAAGAGCGCTCAAGAAAGTACGATATTACT	240
Db	300	TGTCCTTATGTAAAGATATTAACCAAAAGAGCGCTCAAGAAAGTACGATATTACT	359
OY	241	CAACTTGTGTAAGACCTATTGAAAATCATTTTGCTTTCAGCTGTACACAGCTTTGAG	300
Db	360	CAACTTGTGTAAGACCTATTGAAAATCATTTTGCTTTCAGCTGTACACAGCTTTGAG	419
OY	301	TATGCAAAACGCTATTAATTTTGC AAAAAGAAATTAACCTTCCTGAGACCTATAAAGAT	360
Db	420	TATGCAAAACGCTATTAATTTTGC AAAAAGAAATTAACCTTCCTGAGACCTATAAAGAT	479
OY	361	CAAGCTTTATCATTCCAAAAGTATGGGCTACAGAAACCGTGCCAAAAGACTTCTACAGAGT	420
Db	480	GAAAGTTTCTATCATTCCAAAAGTATGGGCTACAGAAACCGTGCCAAAAGACTTCTACAGAGT	539
OY	421	GAACCCGAAAATCCTTCCTGCGAGAAACCACTGCAGTCCCACTGCTTAACCTTGGGA	480
Db	540	GAACCCGAAAATCCTTCCTGCGAGAAACCACTGCAGTCCCACTGCTTAACCTTGGGA	599
OY	481	ACTGTGAGAACTCTGAGGACAAGGACGGGATACAACTCAAAAAGACGTCTGCTACATT	540
Db	600	ACTGTGAGAACTCTGAGGACAAGGACGGGATACAACTCAAAAAGACGTCTGCTACATT	659
OY	541	GAAATGGGATCTGATTTCTTCTGAGAAATACCGTTAATTAAGGCAACTTAATTCAGTGTGGGA	600
Db	660	GAAATGGGATCTGATTTCTTCTGAGAAATACCGTTAATTAAGGCAACTTAATTCAGTGTGGGA	719
OY	601	GATCAGAAATGTTCAAATCAACCCCTCAAGGAAACGAGGATGAATACTAGTTGGATTCT	660
Db	720	GATCAGAAATGTTCAAATCAACCCCTCAAGGAAACGAGGATGAATACTAGTTGGATTCT	779
OY	661	GCAAAAAAGCGCTTGTGAATTTTCTGAGAGGATGTAACAAAATCTAACAATCTATCA	720
Db	780	GCAAAAAAGCGCTTGTGAATTTTCTGAGAGGATGTAACAAAATCTAACAATCTATCA	839

QY 721 CCCAGTAAATGATTTTGAACACACCTGAGAAAGCGTGACGTGAGAGCATCCAGAAAG 780  
|||||  
Db 840 CCCAGTAATAATGATTTTGAACACACCTGAGAAAGCGTGACGTGAGAGCATCCAGAAAG 899  
QY 781 TATCAGGGTAGTTCTTTTCAAACTTGATGGAGCCATGTGGCACAATTAATCTCATGCC 840  
|||||  
Db 900 TATCAGGGTAGTTCTTTTCAAACTTGATGGAGCCATGTGGCACAATTAATCTCATGCC 959  
QY 841 ACCTCATTTACAGATGAGAAAGACAGCATTTTATTTACTCAGTAAAGACAGAAATGATGAGAA 900  
|||||  
Db 960 ACCTCATTTACAGATGAGAAAGACAGCATTTTATTTACTCAGTAAAGACAGAAATGATGAGAA 1019  
QY 901 AAGCGTGAATTTCTGTATATAAAGCAAAACAGCCCTGGCTTTAGCAAGAGCCCAATTAACAGA 960  
1020 AAGCGTGAATTTCTGTATATAAAGCAAAACAGCCCTGGCTTTAGCAAGAGCCCAATTAACAGA 1079  
QY 961 TGGGCTGGAGTAGTAAGAAACATGTATGATAGCGGAGCTCCGACACAGAAAAAAGTA 1020  
|||||  
Db 1080 TGGGCTGGAGTAGTAAGAAACATGTATGATAGCGGAGCTCCGACACAGAAAAAAGTA 1139  
QY 1021 GATCTGAATGCTGATCCCTGTGTGAGAGAAAAAGTAAGAAAGCAAGCACTGCGCATGC 1080  
|||||  
Db 1140 GATCTGAATGCTGATCCCTGTGTGAGAGAAAAAGTAAGAAAGCAAGCACTGCGCATGC 1199  
QY 1081 TCAGAGAAATCTAGAGATCTGAAAGATGTTCTTGGATAACACTAAATAGCAGCATTCAG 1140  
|||||  
Db 1200 TCAGAGAAATCTAGAGATCTGAAAGATGTTCTTGGATAACACTAAATAGCAGCATTCAG 1259  
QY 1141 AAGGTTAATGATGGTTTCCAGAAATGAACTGTAGGTTCTGTATGATGATCTCAGATGAT 1200  
|||||  
Db 1260 AAGGTTAATGATGGTTTCCAGAAATGAACTGTAGGTTCTGTATGATGATCTCAGATGAT 1319  
QY 1201 GGGGAGTCTGAATCAAAATGCCAAAGTAGTGATGTTGGACGTTCTAAATGAGTAGAT 1260  
|||||  
Db 1320 GGGGAGTCTGAATCAAAATGCCAAAGTAGTGATGTTGGACGTTCTAAATGAGTAGAT 1379  
QY 1261 GAATATTTCTGGTTCTTCAAGAAATAGACTTAAGTGGCCAGTATCCTCATGAGGCTTA 1320  
|||||  
Db 1380 GAATATTTCTGGTTCTTCAAGAAATAGACTTAAGTGGCCAGTATCCTCATGAGGCTTA 1439  
QY 1321 AATATGAAAGTGAAGAGTTCACCTCAATCAGTAGAGAGTAATTTGAAGACAAATA 1380  
|||||  
Db 1440 AATATGAAAGTGAAGAGTTCACCTCAATCAGTAGAGAGTAATTTGAAGACAAATA 1499  
QY 1381 TTTGGGAAAACCTATCGAAGAGCAAGCCCTCCCACTTAAGCCATGTAAGTGAAT 1440  
|||||  
Db 1500 TTTGGGAAAACCTATCGAAGAGCAAGCCCTCCCACTTAAGCCATGTAAGTGAAT 1559  
QY 1441 CTAATTTATAGAGCATTTGTTACTGAGCCACAGATTAATACAGAGCGTCCCTCACAAT 1500  
|||||  
Db 1560 CTAATTTATAGAGCATTTGTTACTGAGCCACAGATTAATACAGAGCGTCCCTCACAAT 1619  
QY 1501 AAATTAAGCCATAAAGGAGCTCATCAGGCTTCATCCAGGATTTTATCAAGAAA 1560  
|||||  
Db 1620 AAATTAAGCCATAAAGGAGCTCATCAGGCTTCATCCAGGATTTTATCAAGAAA 1679  
QY 1561 GCAGATTTGGCAGTTCAAAAAGACTCCTGAATGATTAATCAGGAGACTAAACAAAGGAG 1620  
1680 GCAGATTTGGCAGTTCAAAAAGACTCCTGAATGATTAATCAGGAGACTAAACAAAGGAG 1739  
QY 1621 CAGAAATGGTCAAGTATGATTAATTAATTAATGATGATGATGATGATGATGATGATGAT 1680  
|||||  
Db 1740 CAGAAATGGTCAAGTATGATTAATTAATTAATGATGATGATGATGATGATGATGATGAT 1799  
QY 1681 TCTATTCAGAAATGAGAAAAATCCTAACCAATAGATTCATCGAAAAAGAAATGCTTTC 1740  
|||||  
Db 1800 TCTATTCAGAAATGAGAAAAATCCTAACCAATAGATTCATCGAAAAAGAAATGCTTTC 1859  
QY 1741 AAAAGGAAAGCTGAACCTATTAAGCAGCAGTATTAAGCAATATGGAATCGAATTAATATTC 1800  
|||||  
Db 1860 AAAAGGAAAGCTGAACCTATTAAGCAGCAGTATTAAGCAATATGGAATCGAATTAATATTC 1919  
QY 1801 CACAATTCAAAAAGCACTTAATAAAGATAGGCTGAGAGGAGAGTCTTCTACAGGATATTT 1860  
|||||

Db 1920 CACAATTCAAAAAGCACTTAATAAAGATAGGCTGAGAGGAGAGTCTTCTACAGGATATTT 1979  
QY 1861 CATGCCCTGAACTACTACTAGTACGATAGAAATCTTAAGCCCACTTAATTTGATCGAATTTGAA 1920  
|||||  
Db 1980 CATGCCCTGAACTACTACTAGTACGATAGAAATCTTAAGCCCACTTAATTTGATCGAATTTGAA 2039  
QY 1921 ATTGAATGTTGTTCTGACAGTGAAGATTAAGAAAAAAGTATACACCAATATGCCACTC 1980  
|||||  
Db 2040 ATTGAATGTTGTTCTGACAGTGAAGATTAAGAAAAAAGTATACACCAATATGCCACTC 2099  
QY 1981 AGGCACAGAGAAACCTTACACTCATGAGGTAAGAACTGCAATCGAGCCCAAGAG 2040  
2100 AGGCACAGAGAAACCTTACACTCATGAGGTAAGAACTGCAATCGAGCCCAAGAG 2159  
QY 2041 AGTAACACCCAAATGATAACAGCAAGTAAACATGACAGTATGATCTTCCAGAGCTG 2100  
|||||  
Db 2160 AGTAACACCCAAATGATAACAGCAAGTAAACATGACAGTATGATCTTCCAGAGCTG 2219  
QY 2101 AAGTTAACAAATGACACCTGTTCTTTTACTAGTGTCAATATACAGTGAACCTTAAGAA 2160  
|||||  
Db 2220 AAGTTAACAAATGACACCTGTTCTTTTACTAGTGTCAATATACAGTGAACCTTAAGAA 2279  
QY 2161 TTTGTCAATCTAGCCTTCCAGAGAGAAAAAAGAGAAACTAGAAACAGTTAAAGTG 2220  
|||||  
Db 2280 TTTGTCAATCTAGCCTTCCAGAGAGAAAAAAGAGAAACTAGAAACAGTTAAAGTG 2339  
QY 2221 TCTAATTAATGCTGAACCCCAAGATCTCATGTTAAGTGGGAAAGGTTTTCGAAACT 2280  
|||||  
Db 2340 TCTAATTAATGCTGAACCCCAAGATCTCATGTTAAGTGGGAAAGGTTTTCGAAACT 2399  
QY 2281 GAAAGATCTGTAGAGAGTAGCAGTATTTCAATTTGTAAGTCTGATGATTAAGGAGCAGCTCAG 2340  
|||||  
Db 2400 GAAAGATCTGTAGAGAGTAGCAGTATTTCAATTTGTAAGTCTGATGATTAAGGAGCAGCTCAG 2459  
QY 2341 GAAAGATCTGTAGAGAGTAGCAGTATTTGAAAGAGGAGGAGCTAATTCATGTTGTTCCAAA 2400  
|||||  
Db 2460 GAAAGATCTGTAGAGAGTAGCAGTATTTGAAAGAGGAGGAGCTAATTCATGTTGTTCCAAA 2519  
QY 2401 TGTGTAGTCAAGTGTGACAGTATTTGAAAGAGGAGGAGCTAATTCATGTTGTTCCAAA 2460  
|||||  
Db 2520 TGTGTAGTCAAGTGTGACAGTATTTGAAAGAGGAGGAGCTAATTCATGTTGTTCCAAA 2579  
QY 2461 GATTAATGAAATGACACAGAGGCTTTAAGTATTCATTTGGAGCATGAATTAACACAGT 2520  
2580 GATTAATGAAATGACACAGAGGCTTTAAGTATTCATTTGGAGCATGAATTAACACAGT 2639  
QY 2521 CGGGAACAAAGCATTAAGAAATGGAAGAAAGTGAATGATGCTCAGTATTTGACAGATACA 2580  
|||||  
Db 2640 CGGGAACAAAGCATTAAGAAATGGAAGAAAGTGAATGATGCTCAGTATTTGACAGATACA 2699  
QY 2581 TTTCAAGGTTTCAAAAGCGCCAGTATTTGCTGTTTCAAAATCGAGAAATGACAAAG 2640  
2700 TTTCAAGGTTTCAAAAGCGCCAGTATTTGCTGTTTCAAAATCGAGAAATGACAAAG 2759  
QY 2641 GAATGTGCAAACTTCTGCCCCACTGTGGGCTTTAAGAAAAAAGTCCAAAGTCACT 2700  
2760 GAATGTGCAAACTTCTGCCCCACTGTGGGCTTTAAGAAAAAAGTCCAAAGTCACT 2819  
QY 2701 TTTGAATGTGAACAAAGGAGAAAGAAATCAAGAAAGATGAGTCTTAATATCAAGCTGTA 2760  
|||||  
Db 2820 TTTGAATGTGAACAAAGGAGAAAGAAATCAAGAAAGATGAGTCTTAATATCAAGCTGTA 2879  
QY 2761 CAGACAGTTAATATCACTGACAGCTTTCCTGTGGTGTGAGAAAGATTAACAGTTCAT 2820  
|||||  
Db 2880 CAGACAGTTAATATCACTGACAGCTTTCCTGTGGTGTGAGAAAGATTAACAGTTCAT 2939  
QY 2821 AATGCCAAATGTAGTATCAAGAGAGCTCTAGGTTTGTCTATCATCTCAAGTTCAGAGGC 2880  
|||||  
Db 2940 AATGCCAAATGTAGTATCAAGAGAGCTCTAGGTTTGTCTATCATCTCAAGTTCAGAGGC 2999  
QY 2881 AACGAACGTGACCTATTAATCAATTAACATGACATTTTATCAAAACCCATATCGATTA 2940  
|||||

```
Db 3000 AACGAACCTGACTATCTACCAATTAACATGACCTTTTACAAAACCCATATCGTATA 3059
QY 2941 CCACCACTTTTCCCATCAGATCTGTGTAAACCTAAATGTAGAAGAAATCTGCTAGAG 3000
Db 3060 CCACCACTTTTCCCATCAGATCTGTGTAAACCTAAATGTAGAAGAAATCTGCTAGAG 3119
QY 3001 GAAAACCTTTGAGGAACATTCATGTCACCTGAAAGAGAAATGGGAAATGAGAACATTTCCA 3060
Db 3120 GAAAACCTTTGAGGAACATTCATGTCACCTGAAAGAGAAATGGGAAATGAGAACATTTCCA 3179
QY 3061 AGTACATGAGCACAATTTACCCGTAATACATTAGAGAAATGTTTTTAAAGAACCCAGC 3120
Db 3180 AGTACATGAGCACAATTTACCCGTAATACATTAGAGAAATGTTTTTAAAGAACCCAGC 3239
QY 3121 TCAGCAATATTTATGAGTAGAGTTCAGTACTAATGAGTGGGCTCAGTATTAATGAA 3180
Db 3240 TCAGCAATATTTATGAGTAGAGTTCAGTACTAATGAGTGGGCTCAGTATTAATGAA 3299
QY 3181 ATAGGTTCCAGTATGAGTAAACATTTCAAGCAGAACTAGTAGAAAACAGAGGGCCAAAATG 3240
Db 3300 ATAGGTTCCAGTATGAGTAAACATTTCAAGCAGAACTAGTAGAAAACAGAGGGCCAAAATG 3359
QY 3241 AATGCTATGCTTATGATTTAGAGGGTTTTCACACCTGAGGCTATATAACAAATCTTCTGGA 3300
Db 3360 AATGCTATGCTTATGATTTAGAGGGTTTTCACACCTGAGGCTATATAACAAATCTTCTGGA 3419
QY 3301 AGTATATGTAAGCATCTGTAATTAATAAAGCAAGAAATATGAGAGTAGTTCAGACTGT 3360
Db 3420 AGTATATGTAAGCATCTGTAATTAATAAAGCAAGAAATATGAGAGTAGTTCAGACTGT 3479
QY 3361 AATACAGATTTCTCTCCATATCTGATTTTCAGATTAAGTACAGACGCTATGGGAGTAGT 3420
Db 3480 AATACAGATTTCTCTCCATATCTGATTTTCAGATTAAGTACAGACGCTATGGGAGTAGT 3539
QY 3421 CATGCACTCTAGGTTTGTGTGAGACACCTGATGAGCTGTTAGATGATGTTGAATTAAG 3480
Db 3540 CATGCACTCTAGGTTTGTGTGAGACACCTGATGAGCTGTTAGATGATGTTGAATTAAG 3599
QY 3481 GAAGATCTAGTTTGTGTGAATAATGACATTAAAGAAAGTCTGCTGTTTACAAAAGC 3540
Db 3600 GAAGATCTAGTTTGTGTGAATAATGACATTAAAGAAAGTCTGCTGTTTACAAAAGC 3659
QY 3541 GTCCAGAAAGAGAGCTTTAGCAGAGAGTCTAGCCCTTTCACCCATACATTTGGCTCAG 3600
Db 3660 GTCCAGAAAGAGAGCTTTAGCAGAGAGTCTAGCCCTTTCACCCATACATTTGGCTCAG 3719
QY 3601 GGTACCGAAGAGAGGCGCAAGAAATTAAGTCTCAGAGAGAACTTATCTAGTAGAT 3660
Db 3720 GGTACCGAAGAGAGGCGCGAAGAAATTAAGTCTCAGAGAGAACTTATCTAGTAGAT 3779
QY 3661 GAAGAGCTTCCCTGCTTCCAAACCTGTTATTTGGTAAAGTAAACAATATACCTTCAG 3720
Db 3780 GAAGAGCTTCCCTGCTTCCAAACCTGTTATTTGGTAAAGTAAACAATATACCTTCAG 3839
QY 3721 TCTACTAGGATAGACACCGTGTGTACGAGTGTCTGTCTAAGAAACAGAGAGAAATTA 3780
Db 3840 TCTACTAGGATAGACACCGTGTGTACGAGTGTCTGTCTAAGAAACAGAGAGAAATTA 3899
QY 3781 TTATCATTTGAAGAAATAGCTTAATAGCTGAGTACAGAGTAAATTTGGCAAGGATCT 3840
Db 3900 TTATCATTTGAAGAAATAGCTTAATAGCTGAGTACAGAGTAAATTTGGCAAGGATCT 3959
QY 3841 CAGGACATCACCTTGTGTGAGAAACAAATGTTGCTAGTGTGTTTCTTACAGTGC 3900
Db 3960 CAGGACATCACCTTGTGTGAGAAACAAATGTTGCTAGTGTGTTTCTTACAGTGC 4019
QY 3901 AGTAAATGGAAGACTTGTAGTCAAAATACAAACACCCAGAGATCTTCTTATGATGTTCT 3960
Db 4020 AGTAAATGGAAGACTTGTAGTCAAAATACAAACACCCAGAGATCTTCTTATGATGTTCT 4079
QY 3961 TCCAAACAAATGAGCATGAGTCTGAAGCCAGGAGGAGTGTGCTGAGTCAAGAAATG 4020
Db 4080 TCCAAACAAATGAGCATGAGTCTGAAGCCAGGAGGAGTGTGCTGAGTCAAGAAATG 4139
QY 4021 GTTTCAGATGATGAGAAAGAGAACCGGCTTGGAAAGAAATATATCAAGAGACAAAGC 4080
Db 4140 GTTTCAGATGATGAGAAAGAGAACCGGCTTGGAAAGAAATATATCAAGAGACAAAGC 4199
QY 4081 ATGGATTCAAACTTATAGTGAAGCAGCATCTGGGTGTGAGAGTGAACAGCGTCTGTAA 4140
Db 4200 ATGGATTCAAACTTATAGTGAAGCAGCATCTGGGTGTGAGAGTGAACAGCGTCTGTAA 4259
QY 4141 GACTGCTCAGGGCTATCTCTCAGAGTGAATTTTAAACCTCAGCAGAGAGGATACATG 4200
Db 4260 GACTGCTCAGGGCTATCTCTCAGAGTGAATTTTAAACCTCAGCAGAGAGGATACATG 4319
QY 4201 CAACATTAACCTGATTAAGGTCAGCAGAGAAATGGCTCAATAGAGCTGTTTGAAGAC 4260
Db 4320 CAACATTAACCTGATTAAGGTCAGCAGAGAAATGGCTCAATAGAGCTGTTTGAAGAC 4379
QY 4261 CATGGAGCAGCCCTTCTAACAGCTACCCCTTCATCATAGTACCTTCTGCTTGGAG 4320
Db 4380 CATGGAGCAGCCCTTCTAACAGCTACCCCTTCATCATAGTACCTTCTGCTTGGAG 4439
QY 4321 GACCTGGAAATCCAGAACAAAGCAGATCCAGAAAGCCCTTCTGTCAGCAAGTTGAGTGTCT 4440
Db 4440 GACCTGGAAATCCAGAACAAAGCAGATCCAGAAAGCCCTTCTGTCAGCAAGTTGAGTGTCT 4499
QY 4381 AGTGAATACCTTATTAAGCCAGAAATCCAGAAAGCCCTTCTGTCAGCAAGTTGAGTGTCT 4440
Db 4500 AGTGAATACCTTATTAAGCCAGAAATCCAGAAAGCCCTTCTGTCAGCAAGTTGAGTGTCT 4559
QY 4441 GCAGATAGTTCTACCACTAATAAATAAAGAACACAGAGTGAAGAGTCATCCCTCTTAAA 4500
Db 4560 GCAGATAGTTCTACCACTAATAAATAAAGAACACAGAGTGAAGAGTCATCCCTCTTAAA 4619
QY 4501 TGGCCATCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4560
Db 4620 TGGCCATCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4679
QY 4561 AACTACCATCTCAAGAGAGAGCTCATTAAGTGTGATGATGATGATGATGATGATGATGAT 4620
Db 4680 AACTACCATCTCAAGAGAGAGCTCATTAAGTGTGATGATGATGATGATGATGATGATGAT 4739
QY 4621 GAGTCTGGGCGACAGCATTTTGAAGGAAACATCTTACTTGGCAAGGCAAGATCTAGAGGA 4680
Db 4740 GAGTCTGGGCGACAGCATTTTGAAGGAAACATCTTACTTGGCAAGGCAAGATCTAGAGGA 4799
QY 4681 ACCCTTACCTGGAATCTGGAATCAGAGCTCTCTCTGATGATGATGATGATGATGATGAT 4740
Db 4800 ACCCTTACCTGGAATCTGGAATCAGAGCTCTCTCTGATGATGATGATGATGATGATGAT 4859
QY 4741 GAAGACAGAGCCCGAGAGTACGTCGTTGGCAACATACATCTTACCTGTCATG 4800
Db 4860 GAAGACAGAGCCCGAGAGTACGTCGTTGGCAACATACATCTTACCTGTCATG 4919
QY 4801 AAGATTCGCCAATTTGAAGTTGCAAGATCTGCCAGAGTCCAGCTGCTCATACTACT 4860
Db 4920 AAGATTCGCCAATTTGAAGTTGCAAGATCTGCCAGAGTCCAGCTGCTCATACTACT 4979
QY 4861 GATACGCTGGGTATATGCAATGGAAGAAAGTGTGAGCGAGGAGAAAGCCAGAAATGACA 4920
Db 4980 GATACGCTGGGTATATGCAATGGAAGAAAGTGTGAGCGAGGAGAAAGCCAGAAATGACA 5039
QY 4921 GCTTCAACAGAAAGGCTCAACAAAGATGCTCATGATGATGATGATGATGATGATGATGAT 4980
Db 5040 GCTTCAACAGAAAGGCTCAACAAAGATGCTCATGATGATGATGATGATGATGATGATGAT 5099
QY 4981 GAATTTATGCTGCTGTACAGATTTGCGAGAAACACACATCACTTAATTAATTAAT 5040
Db 5100 GAATTTATGCTGCTGTACAGATTTGCGAGAAACACACATCACTTAATTAATTAAT 5159
QY 5041 ACTGAAGAGACTCTATGTTGTTATGAAGAAAGATGCTGATGTTGTTGTTGTTGTTGTTGTT 5100
Db 5160 ACTGAAGAGACTCTATGTTGTTATGAAGAAAGATGCTGATGTTGTTGTTGTTGTTGTTGTT 5219
```



QY	5101	TTGAAATATTTTCAGGAATTCGGGAGGAAATGGGAGTAGTACTATTTCTGGGTACC	5160
Db	5220	CTGAATATTTTCTTAGGAAATGCGGGAGGAAATGGGAGTAGTACTATTTCTGGGTACC	5279
QY	5161	CAGTCATTTTAAAGAAAGAAAATGCTCGAATGAGCATGTTTGAAGTCAGAGAGATGTG	5220
Db	5280	CAGTCATTTTAAAGAAAGAAAATGCTCGAATGAGCATGTTTGAAGTCAGAGAGATGTG	5339
QY	5221	GTCATATGGAACAAACCCAGAGTCGCAAGGAGCACAAGAGATATCCACGAGACAGAAATC	5280
Db	5340	GTCATATGGAABAAACCCAGAGTCGCAAGGAGCACAAGAGATATCCACGAGACAGAAATC	5399
QY	5381	TTTCAGGGGGCTAGAAATCTGTTCCTATGAGGCCCTTACCAACATGCCACAGATCAACTG	5340
Db	5400	TTTCAGGGGGCTAGAAATCTGTTCCTATGAGGCCCTTACCAACATGCCACAGATCAACTG	5459
QY	5341	GAATGGATGGTACAGCTGTGTGGTCTCTGTGTGGTGAAGAGCTTTATATCATTTACCCCTT	5400
Db	5460	GAATGGATGGTACAGCTGTGTGGTCTCTGTGTGGTGAAGAGCTTTATATCATTTACCCCTT	5519
QY	5401	GGCACAGGTGTCCACCCCAATGTGTGTTGTGCAGCCAGATGCCCTGGACAGAGAGCAATGGC	5460
Db	5520	GGCACAGGTGTCCACCCCAATGTGTGTTGTGCAGCCAGATGCCCTGGACAGAGAGCAATGGC	5579
QY	5461	TTTCATATCAATTTGGGCAGATGTGTGAGGCACCTGTGTGACCCAGAGATGGGTCTTGGAC	5520
Db	5580	TTTCATATCAATTTGGGCAGATGTGTGAGGCACCTGTGTGACCCAGAGATGGGTCTTGGAC	5639
QY	5521	AGTTAGACACTCTACCAAGTGCACAGAGCTGACACCTACCTGATACCCACAGATCCCCAC	5580
Db	5640	AGTTAGACACTCTACCAAGTGCACAGAGCTGACACCTACCTGATACCCACAGATCCCCAC	5699
QY	5581	AGGCACATAC 5589	
Db	5700	AGGCACATAC 5708	

XX	RESULT 9
XX	AAV46460
ID	AAV46460 standard; cDNA; 5711 bp.
XX	
AC	AAV46460;
XX	
DT	18-NOV-1998 (first entry)
XX	
DE	Human BRCA1 omi2 polymorphism #3 cDNA.
XX	
KW	BRCA1; omi2; human; breast and ovarian cancer predisposing gene;
KW	polymorphism; susceptibility; anti-oncogene; tumour suppressor;
KM	chromosome 17q; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	120..5711
FT	/tag= a
FT	/product= "BRCA1 omi2 protein"
FT	2731
FT	/tag= b
FT	/note= "This polymorphic variation can be a C or T nucleotide"
PX	
PN	US5750400-A.
XX	
PD	12-MAY-1998.
XX	
PE	12-FEB-1997; 97US-0798691.
XX	
PR	12-FEB-1996; 96US-0598591.
PR	12-FEB-1997; 97US-0798661.
PA	(ONCO-) ONCORMED INC.
XX	

PI	Allen A, Alvares CP, Critz BS, Murphy PD, Olson SJ,
PI	Schelter DB, Zeng B;
DR	WPI; 1998-296774/26.
XX	
XX	BRCA1 om1 gene coding sequences - useful for distinguishing between
PT	polymorphisms and mutation(s) in the screening for disposition to
PT	breast or ovarian cancer
XX	
PS	Claim 2e; Page -; 54pp; English.
XX	
CC	This sequence encodes a human BRCA1 (breast and ovarian cancer
CC	predisposing gene) om12 gene in which a polymorphic variation occurs at
CC	nucleotide 2731. This sequence and other polymorphic variations of this
CC	sequence are useful for the identification of an individual who may or
CC	may not have an increased susceptibility to breast or ovarian cancer.
CC	The sequences used identify gene changes which are due to polymorphisms,
CC	rather than cancer causing mutations. BRCA1 is an anti-oncogene (tumour
CC	suppressor) which is involved in genetic inheritance of cancers,
CC	especially breast and ovarian cancer. It is found at human chromosome
CC	17q which is known to be linked to cancer susceptibility, especially
CC	breast cancer. Cells containing a mutation in this gene lose the
CC	wild-type function of BRCA1 and are more susceptible to cancers.
CC	NOTE: This sequence does not appear in the specification but has been
CC	created from the wild type BRCA1 om12 gene represented in AAV6449.
XX	
SO	Sequence 5711 BP, 1956 A, 1098 C, 1274 G, 1382 T, 1 other;

Query Match	100.0%;	Score 5587;	DB 19;	length 5711;
Best Local Similarity	100.0%;	Pred. NO. 0;		
Matches 5587; Conservative	1;	Mismatches	1;	Indels 0; Gaps 0;

QY	1	ATGATATTATTCGCTTCCTTCGCTTGAAGATACAAAGTCAATATGCTATGACAGAA	60
Db	120	ATGATATTATTCGCTTCCTTCGCTTGAAGATACAAAGTCAATATGCTATGACAGAA	179
QY	61	ATCTTAGAGTGTCCATCTGTCTGAGATTCAGGAACCTGTCTCCACAAAGTGTAC	120
Db	180	ATCTTAGAGTGTCCATCTGTCTGAGATTCAGGAACCTGTCTCCACAAAGTGTAC	239
QY	121	CACATATTTTTCGAATTTTTCGATTCGTGAACCTCTCAACCCAGGAAGAGGCTTCACAG	180
Db	240	CACATATTTTTCGAATTTTTCGATTCGTGAACCTCTCAACCCAGGAAGAGGCTTCACAG	299
QY	181	TGTCCCTTATGTAGACATGATATTAACCCAAAGAGAGCCACAGAAATAGAGATTTAGT	240
Db	300	TGTCCCTTATGTAGACATGATATTAACCCAAAGAGAGCCACAGAAATAGAGATTTAGT	359
QY	241	CAACTGTGTGAAGAGCTATTTGAAATTCATTTGTCTCTTCAGCTTGCACACAGTTTGAG	300
Db	360	CAACTGTGTGAAGAGCTATTTGAAATTCATTTGTCTCTTCAGCTTGCACACAGTTTGAG	419
QY	301	TATGCCAAACAGCTATTAATTTGCAAAAAGGAAATTAACCTCTCTGAAACCTTAAGAAAT	360
Db	420	TATGCCAAACAGCTATTAATTTGCAAAAAGGAAATTAACCTCTCTGAAACCTTAAGAAAT	479
QY	361	GAGGTTTCTATCATGCAAAAGTATGGGTAACAGAAACCGTGCAAAACATCTTACAGAGT	420
Db	480	GAGGTTTCTATCATGCAAAAGTATGGGTAACAGAAACCGTGCAAAACATCTTACAGAGT	539
QY	421	GAAACCCGAAATTCCTCTTTCGAGGAACACAGTCTCAGTGTCCAACTCTCTAACCTTGA	480
Db	540	GAAACCCGAAATTCCTCTTTCGAGGAACACAGTCTCAGTGTCCAACTCTCTAACCTTGA	599
QY	481	ACTGTGAGAACTCTGAGAGCAAAAGCAGGGGATACAACCTCAAAAAGACGTCTGTCTACATT	540
Db	600	ACTGTGAGAACTCTGAGAGCAAAAGCAGGGGATACAACCTCAAAAAGACGTCTGTCTACATT	659
QY	541	GATTTGGAGTCTGATCTCTTTCGAAGATTCACCTTAATTAAGCAACTTTTGCATGTGGGA	600
Db	660	GATTTGGAGTCTGATCTCTTTCGAAGATTCACCTTAATTAAGCAACTTTTGCATGTGGGA	719
QY	601	GATCAAGAAATTTTCACAATCACCCCTCAAGAACCAAGGATGAATCACTTTGGATTCT	660



|||||  
Db 720 GATCAGAAATGTTTACAAATCACCCCTCAAGAGAACGAGGATGAATTCAGTTTGATTTCT 779  
QY 661 GCAAAAAAGGCTCTGTGTAATTTTCTGAGACGAGATGTAACAAATCTGAACATCATCAA 720  
Db 780 GCAAAAAAGGCTCTGTGTAATTTTCTGAGACGAGATGTAACAAATCTGAACATCATCAA 839  
QY 721 CCCAGTAATTAATGATTGTAACACACCTGAGAGACGTCAGACGTGAGAGGATCCAGAAAG 780  
Db 840 CCCAGTAATTAATGATTGTAACACACCTGAGAGACGTCAGACGTGAGAGGATCCAGAAAG 899  
QY 781 TATCAGGAGTATCTCTTTCAAACTTGATGTGAGGACATGTGCACAAAATCTCATGCC 840  
Db 900 TATCAGGAGTATCTCTTTCAAACTTGATGTGAGGACATGTGCACAAAATCTCATGCC 959  
QY 841 AGCTCATATACAGATGAGAACACAGCTTTTATCTACTCATTAACACAGATGATGTA 900  
Db 960 AGCTCATATACAGATGAGAACACAGCTTTTATCTACTCATTAACACAGATGATGTA 1019  
QY 901 AAGGCTGAATTTCTGTATATAAAGCAACAGCTGGCTTACAGAGGCCAATTAACAGA 960  
Db 1020 AAGGCTGAATTTCTGTATATAAAGCAACAGCTGGCTTACAGAGGCCAATTAACAGA 1079  
QY 961 TGGGCTGGAAGTAAAGAAACATGTAATGATAGCGGACCTCCAGACAGAAAAAGGTA 1020  
Db 1080 TGGGCTGGAAGTAAAGAAACATGTAATGATAGCGGACCTCCAGACAGAAAAAGGTA 1139  
QY 1021 GATCTGAATGCTGATCCCTGCTGTGAGAGAAAAAGATGATAGCAAGAACTGCCATGC 1080  
Db 1140 GATCTGAATGCTGATCCCTGCTGTGAGAGAAAAAGATGATAGCAAGAACTGCCATGC 1199  
QY 1081 TCAGAGAAATCTGTAGATGATGTAAGATGTTCTGTGATTAACACTAAATAGCAGATTGCA 1140  
Db 1200 TCAGAGAAATCTGTAGATGATGTAAGATGTTCTGTGATTAACACTAAATAGCAGATTGCA 1259  
QY 1141 AAGGTTAATGAGTGTGTTTTCCAGAGTGAACCTGTTAGCTTGTATGATCAGATGAT 1200  
Db 1260 AAGGTTAATGAGTGTGTTTTCCAGAGTGAACCTGTTAGCTTGTATGATCAGATGAT 1319  
QY 1201 GGGGAGTCTGAAATGCAAAATGCCAAAGTATGATGATGATGATGATGATGATGATGAT 1260  
Db 1320 GGGGAGTCTGAAATGCAAAATGCCAAAGTATGATGATGATGATGATGATGATGATGAT 1379  
QY 1261 GAATATTTGCTGTTCTTCAAGAGAAATAGACTTACTGCGCAGTATCTCTGAGGCTTGA 1320  
Db 1380 GAATATTTGCTGTTCTTCAAGAGAAATAGACTTACTGCGCAGTATCTCTGAGGCTTGA 1439  
QY 1321 ATATGTAAGTGAAGAGTTCACCTCAATCAGTAGAGATTAATTTGAAGCAAAATA 1380  
Db 1440 ATATGTAAGTGAAGAGTTCACCTCAATCAGTAGAGATTAATTTGAAGCAAAATA 1499  
QY 1381 TTTTGGAAAAACCTATGCGAAGAGGCAAGCCCTCCCACTTAAGCCATGTAATGAAAT 1440  
Db 1500 TTTTGGAAAAACCTATGCGAAGAGGCAAGCCCTCCCACTTAAGCCATGTAATGAAAT 1559  
QY 1441 CTATATATAGAGACATTTTGTACTGAGCCACAGATATATACAAAGGCTCCCTCACAAT 1500  
Db 1560 CTATATATAGAGACATTTTGTACTGAGCCACAGATATATACAAAGGCTCCCTCACAAT 1619  
QY 1501 AAATTAAGCGTAAAGAGAGACCTATCAGGCTTCATCTGAGAGATTTTATCAAGAAA 1560  
Db 1620 AAATTAAGCGTAAAGAGAGACCTATCAGGCTTCATCTGAGAGATTTTATCAAGAAA 1679  
QY 1561 GAGATTTTGGAGTTCAAAAGACTCCTGAAATGATTAATCAGGAGACTAACCAAGGAG 1620  
Db 1680 GAGATTTTGGAGTTCAAAAGACTCCTGAAATGATTAATCAGGAGACTAACCAAGGAG 1739  
QY 1621 CAGAAATGCTCAAGTATGATATTTACTAATAGTGTGATGAGAAATTAACAAAGGCTGAT 1680  
Db 1740 CAGAAATGCTCAAGTATGATATTTACTAATAGTGTGATGAGAAATTAACAAAGGCTGAT 1799  
QY 1681 TCTATTCAGATGAGAAAAATCTTAACCAATAGAAATCACTGAAAAAGATCTGCTTTC 1740  
|||||

Db 1800 TCTATTCAGATGAGAAAAATCTTAACCAATAGAAATCACTGAAAAAGATCTGCTTTC 1859  
QY 1741 AAAACGAAACCTGAACTATTAAGACAGATTAATTAAGCAATTAATGAACTGCAATTAATATC 1800  
Db 1860 AAAACGAAACCTGAACTATTAAGACAGATTAATTAAGCAATTAATGAACTGCAATTAATATC 1919  
QY 1801 CACAATTCAAAGACCTTAATAAAGAAATAGGCTGAGAGAGAACTTCTTACAGGACATATT 1860  
Db 1920 CACAATTCAAAGACCTTAATAAAGAAATAGGCTGAGAGAGAACTTCTTACAGGACATATT 1979  
QY 1861 CATGCGCTTGAAGTATGATGATGAGAAATCTAAGCCACCTAATTTGATGTAATGCA 1920  
Db 1980 CATGCGCTTGAAGTATGATGATGAGAAATCTAAGCCACCTAATTTGATGTAATGCA 2039  
QY 1921 ATTGATAGTGTCTGACATGAGAGATTAAGAAAAAGTAAACCAATGCGCATG 1980  
Db 2040 ATTGATAGTGTCTGACATGAGAGATTAAGAAAAAGTAAACCAATGCGCATG 2099  
QY 1981 AGGCACAGCAGAAACCTACACTCATGGAAGTAAAGAACTGCAACTGAGCCAGAG 2040  
Db 2100 AGGCACAGCAGAAACCTACACTCATGGAAGTAAAGAACTGCAACTGAGCCAGAG 2159  
QY 2041 AGTAACAGCCAAATGAACAGACAGTAAAGACATGACAGTATCTTCCAGAGCTG 2100  
Db 2160 AGTAACAGCCAAATGAACAGACAGTAAAGACATGACAGTATCTTCCAGAGCTG 2219  
QY 2101 AAGTTAAACAAATGACCTGCTTTTCTTACTAAGTGTTCATATACAGTGAATTAAGAA 2160  
Db 2220 AAGTTAAACAAATGACCTGCTTTTCTTACTAAGTGTTCATATACAGTGAATTAAGAA 2279  
QY 2161 TTTGTCATCTGAGCTTCCAGAGAAAGAAAGAAAGAACTGAAAGATTAAGTG 2220  
Db 2280 TTTGTCATCTGAGCTTCCAGAGAAAGAAAGAAAGAAAGAACTGAAAGATTAAGTG 2339  
QY 2221 TCTAATATGCTGAAGACCCCAAGATCTCATGTTAAAGTGAGAAAGGTTTGCACACT 2280  
Db 2340 TCTAATATGCTGAAGACCCCAAGATCTCATGTTAAAGTGAGAAAGGTTTGCACACT 2399  
QY 2281 GAAAGATCTGTAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340  
Db 2400 GAAAGATCTGTAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2459  
QY 2341 GAAAGATCTGTAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2400  
Db 2460 GAAAGATCTGTAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2519  
QY 2401 TGTGTAGTCAAGTGTGACGATTTGAAACCCCAAGGAGCTAATTCATGCTGTTCCAAA 2460  
Db 2520 TGTGTAGTCAAGTGTGACGATTTGAAACCCCAAGGAGCTAATTCATGCTGTTCCAAA 2579  
QY 2461 GATTAATGAATGACACAGAGCTTAAATATCCATTTGGGACATGAAAGTTAACCCACT 2520  
Db 2580 GATTAATGAATGACACAGAGCTTAAATATCCATTTGGGACATGAAAGTTAACCCACT 2639  
QY 2521 CGGGAACAAGCATTAATAATGGAAGAAAGTGAAGTGTGATGATGATGATGATGATGATGAT 2580  
Db 2640 CGGGAACAAGCATTAATAATGGAAGAAAGTGAAGTGTGATGATGATGATGATGATGATGAT 2699  
QY 2581 TTCAGGTTTCAAGCGCCAGTCATTTCTCTGTTTCAATCCAGAAATGAGAGAG 2640  
Db 2700 TTCAGGTTTCAAGCGCCAGTCATTTCTCTGTTTCAATCCAGAAATGAGAGAG 2759  
QY 2641 GAATGTGCAACATTTCTGCGCACCTGCGGTCTTAAAGAAACCAAGGTCACAT 2700  
Db 2760 GAATGTGCAACATTTCTGCGCACCTGCGGTCTTAAAGAAACCAAGGTCACAT 2819  
QY 2701 TTTGAATGTGAACAAAGAGAAATTAAGAAATTAAGTCTAATATCAAGCCCTGTA 2760  
Db 2820 TTTGAATGTGAACAAAGAGAAATTAAGAAATTAAGTCTAATATCAAGCCCTGTA 2879  
QY 2761 CAGACAGTTAATATCACTGAGGCTTCTGCTGTTGTGTCAGAAAGTAAAGCAGTTGAT 2820  
Db 2880 CAGACAGTTAATATCACTGAGGCTTCTGCTGTTGTGTCAGAAAGTAAAGCAGTTGAT 2939  
|||||

QY	2821	AATGCCAAATGTAGTATCAAGAGAGGCTCTAGGTTTGTCTATCATCTCAGCTTCAGAGC	2880
Db	2940	AATGCCAAATGTAGTATCAAGAGAGGAGGCTCTAGGTTTGTCTATCATCTCAGTTCAGAGGC	2999
QY	2881	AACGAAACTGGACCTATTACTCTCCAAATTAACATGACCTTTTACAAACCCATATCGATTA	2940
Db	3000	AACGAAACTGGACCTATTACTCTCCAAATTAACATGACCTTTTACAAACCCATATCGATTA	3059
QY	2941	CCACACATTTTCCCATCAAGTCATTTGTAAAACTAATGTAAAGAAAAATCTGCTAGAG	3000
Db	3060	CCACACATTTTCCCATCAAGTCATTTGTAAAACTAATGTAAAGAAAAATCTGCTAGAG	3119
QY	3001	GAAACTTTGAGGAACTTCATCTGTCACCTGAAAGAGAAATGGGAAATGACAACTTCCA	3060
Db	3120	GAAACTTTGAGGAACTTCATCTGTCACCTGAAAGAGAAATGGGAAAGAACAACTTCCA	3179
QY	3061	AGTACAGTGAGCACAATTAGCCGTATTAACATTAGAGAAATGTTTTAAAGAACCCAGC	3120
Db	3180	AGTACAGTGAGCACAATTAGCCGTATTAACATTAGAGAAATGTTTTAAAGAACCCAGC	3239
QY	3121	TCACCAATATTAAATGAAGTAGGTTCCAGTACTAATGAAGTGGGCTCCAGTATTAAATGA	3180
Db	3240	TCACCAATATTAAATGAAGTAGGTTCCAGTACTAATGAAGTGGGCTCCAGTATTAAATGA	3299
QY	3181	ATAGGTTCCAGTGATGAAGAAACATTCACACAGAACTAGTAAACACAGAGGGCCAAAATTG	3240
Db	3300	ATAGGTTCCAGTGATGAAGAAACATTCACACAGAACTAGTAAACACAGAGGGCCAAAATTG	3359
QY	3241	AATGCTATGCTTAGATTTAGATTTAGGGGTTTGGACCTGAGGCTATTAACAAGCTTCTCGGA	3300
Db	3360	AATGCTATGCTTAGATTTAGATTTAGGGGTTTGGACCTGAGGCTATTAACAAGCTTCTCGGA	3419
QY	3301	AGTAATTGTAAAGCATCTCGAAATTAACAAAGCAGATATGAAGAAATGTTTCAGACTGTT	3360
Db	3420	AGTAATTGTAAAGCATCTCGAAATTAACAAAGCAGATATGAAGAAATGTTTCAGACTGTT	3479
QY	3361	AATACAGATTTCTCTCCATCTGCTGTTCACATTAACATAGAAACAGCCATGCGGAAGTAGT	3420
Db	3480	AATACAGATTTCTCTCCATCTGCTGTTCACATTAACATAGAAACAGCCATGCGGAAGTAGT	3539
QY	3421	CATGCATCTCAGGTTTGTCTGAGACACACTGATGACCGCTGTAGATGATGGGAATTAAG	3480
Db	3540	CATGCATCTCAGGTTTGTCTGAGACACACTGATGACCGCTGTAGATGATGGGAATTAAG	3599
QY	3481	GAAGTACTACTTTTGGCTGAATAATGACATTAAAGAAAGTCTGCTGTTTATAGCAAAAGC	3540
Db	3600	GAAGTACTACTTTTGGCTGAATAATGACATTAAAGAAAGTCTGCTGTTTATAGCAAAAGC	3659
QY	3541	GTCCAGAAAGAGAGCTTAGCAGAGAGGCTGAGCCCTTTCACCCATPACACATTTGGCTCAG	3600
Db	3660	GTCCAGAAAGAGAGCTTAGCAGAGAGGCTGAGCCCTTTCACCCATPACACATTTGGCTCAG	3719
QY	3601	GGTACCCGAAAGGGGCCAAGAAATTTAGAGTCTCAGAAAGAACTTATCTAGTAGAGAT	3660
Db	3720	GGTACCCGAAAGGGGCCAAGAAATTTAGAGTCTCAGAAAGAACTTATCTAGTAGAGAT	3779
QY	3661	GAAAGAGTCTCCCTGCTTCCAACTCTGTTATTTGGTAAAGTAAACAATATACCTTTCAG	3720
Db	3780	GAAAGAGTCTCCCTGCTTCCAACTCTGTTATTTGGTAAAGTAAACAATATACCTTTCAG	3839
QY	3721	TCTACTAGAGCATACCGCTGTCTACGAGTGTCTAGTAAGACACAGAGGAGCAATTTA	3780
Db	3840	TCTACTAGAGCATACCGCTGTCTACGAGTGTCTAGTAAGACACAGAGGAGCAATTTA	3899
QY	3781	TTATCATTTGAAGATAGCTTAAATGACATGCAGTAACAGAGTAATATTGGCAAGGCACTCT	3840
Db	3900	TTATCATTTGAAGATAGCTTAAATGACATGCAGTAACAGAGTAATATTGGCAAGGCACTCT	3959
QY	3841	CAGGAAATCATCCTTAGTAGAGGAAACAAATGTTCTGCTAGCTTGTCTTCCACAGTGC	3900
Db	3960	CAGGAAATCATCCTTAGTAGAGGAAACAAATGTTCTGCTAGCTTGTCTTCCACAGTGC	4019

QY	3901	ACTGAATTGGAGACTGTGACTGCCAATATACAAACACCCAGGATCCCTTCTGTATTTGGTCT	3980
Db	4020	AGTGAATTGGAGACTGTGACTGCCAATATACAAACACCCAGGATCCCTTCTGTATTTGGTCT	4079
QY	3961	TCCAAACAATATAGGCATCATGCTCTGAAGCCAGGGAGTGGCTGTAGTGACAGAATTG	4020
Db	4080	TCCAAACAATATAGGCATCATGCTCTGAAGCCAGGGAGTGGCTGTAGTGACAGAATTG	4139
QY	4021	GTTTCAGATGATGAAGAAAGAGAAACGGGCTTGGAGAAATAATCAAGAAAGCAAGC	4080
Db	4140	GTTTCAGATGATGAAGAAAGAGAAACGGGCTTGGAGAAATAATCAAGAAAGCAAGC	4199
QY	4081	ATGGAATTCAAACTTATAGTGTAAGCACCATCTGGGTGTAGAGTGAAGCAAGCTCTCTAA	4140
Db	4200	ATGGAATTCAAACTTATAGTGTAAGCACCATCTGGGTGTAGAGTGAAGCAAGCTCTCTAA	4239
QY	4141	GACTCTCAGGGCTATTCCTCTCAGAGTGATCATTTTAACCACTCAGCAGAGGATACATG	4200
Db	4260	GACTCTCAGGGCTATTCCTCTCAGAGTGATCATTTTAACCACTCAGCAGAGGATACATG	4319
QY	4201	CAACATTAACCTGATTAAGCTCCAGAGGAATGGCTGAAGCTAGAAAGCTGTATTAGACAG	4260
Db	4320	CAACATTAACCTGATTAAGCTCCAGAGGAATGGCTGAAGCTAGAAAGCTGTATTAGACAG	4379
QY	4261	CATGGGAGCCACCCTTTACAGCGTACCCCTTCATCATTAAGTACTCTTGTGCCCTTAG	4320
Db	4380	CATGGGAGCCACCCTTTACAGCGTACCCCTTCATCATTAAGTACTCTTGTGCCCTTAG	4439
QY	4321	GACTCGCGAATCCAGAACAAAGCAGATCAGAAAGCAGTATTAATCTTACAGAAAGT	4380
Db	4440	GACTCGCGAATCCAGAACAAAGCAGATCAGAAAGCAGTATTAATCTTACAGAAAGT	4499
QY	4381	AGTGAATACCCATATAGCCAGAAATCCAGAAAGGCTTTCTGCTAGCAAGTTTGAGGTGCT	4440
Db	4500	AGTGAATACCCATATAGCCAGAAATCCAGAAAGGCTTTCTGCTAGCAAGTTTGAGGTGCT	4559
QY	4441	GCAGATAGTTCTTACCAGATAAAAATTAAGACAGAGATGGAAAGTCTATCCCTCTTAA	4500
Db	4560	GCAGATAGTTCTTACCAGATAAAAATTAAGACAGAGATGGAAAGTCTATCCCTCTTAA	4619
QY	4501	TGCCCATCATTTAGATGATAGGTGTACATGCACATCTGCTGGGAGCTTCAGAAATGA	4560
Db	4620	TGCCCATCATTTAGATGATAGGTGTACATGCACATCTGCTGGGAGCTTCAGAAATGA	4679
QY	4561	AACATACCATCTCAAGAGAGAGCTCATTTAAGTTGTGTGATGTGAGAGCAACAGCTGGA	4620
Db	4680	AACATACCATCTCAAGAGAGAGCTCATTTAAGTTGTGTGATGTGAGAGCAACAGCTGGA	4739
QY	4621	GAGTTGGGCCACACGATTTTGACGGAACATCTTACTTGCCAGGCAAGATCTAGAGGA	4680
Db	4740	GAGTTGGGCCACACGATTTTGACGGAACATCTTACTTGCCAGGCAAGATCTAGAGGA	4799
QY	4681	ACCCCTTACCTGATCTGGAATCAAGCCCTTCCTGTGATGACCCGGAATCTGATCTCT	4740
Db	4800	ACCCCTTACCTGATCTGGAATCAAGCCCTTCCTGTGTGATGACCCGGAATCTGATCTCT	4859
QY	4741	GAAGACAGAGCCCCAGAGTCAGGCTGTGTGGCAACATACCATCTTCAACCTTCGATTG	4800
Db	4860	GAAGACAGAGCCCCAGAGTCAGGCTGTGTGGCAACATACCATCTTCAACCTTCGATTG	4919
QY	4801	AAAGTTCCCAATTTGAAGTTGACAGATCTGCCAGAGTCCAGGCTGCTCATTAATCT	4860
Db	4920	AAAGTTCCCAATTTGAAGTTGACAGATCTGCCAGAGTCCAGGCTGCTCATTAATCT	4979
QY	4861	GATACTGTGGGTATTAATGCAATGGAGAAGTGTGAGCAGGGGAAGCCAGAAATTGACA	4920
Db	4980	GATACTGTGGGTATTAATGCAATGGAGAAGTGTGAGCAGGGGAAGCCAGAAATTGACA	5039
QY	4921	GCTTCAACAGAAAGGTCACAAAAGATGTGCATGGTGGGTGTGAGGCCCAAGAA	4980
Db	5040	GCTTCAACAGAAAGGTCACAAAAGATGTGCATGGTGGGTGTGAGGCCCAAGAA	5099
QY	4981	GAATTTATGCTGTGTACAGTTTGGCCAGAAACACGACATCACTTTAATTAATCTAAT	5040

```

Db      5100 GAATTTATGCTCGTACAAAGTTGGCCAGAAAACACACATCACTTAATCTAATT 5159
OY      5041 ACAGAGAGACTACTCATGTTTATGAAGAACAGATCTAGATTGTGTGTAACGACA 5100
Db      5160 ACTGAAGAGACTACTCATGTTTATGAAGAACAGATCTAGATTGTGTGTAACGACA 5219
OY      5101 CTGAATATTTTCTAGGAATTTGGGAGAGAAAATGGTAGTTAGCTATTTCTGGGTACC 5160
Db      5220 CTGAATATTTTCTAGGAATTTGGGAGAGAAAATGGTAGTTAGCTATTTCTGGGTACC 5279
OY      5161 CAGTCTATTAAGAAAGAAAATGCTGAATGAGCATGTTTGAAGTCAAGAGAGATGTG 5220
Db      5280 CAGTCTATTAAGAAAGAAAATGCTGAATGAGCATGTTTGAAGTCAAGAGAGATGTG 5339
OY      5221 GTCAATGAGAAACCCCAAGGTCCCAAGCAGAGAAAGATCCAGAGAGAGATG 5280
Db      5340 GTCAATGAGAAACCCCAAGGTCCCAAGCAGAGAAAGATCCAGAGAGAGATG 5399
OY      5281 TTCAGGGGGCTAGAAATCTGTCTATGAGGCCCTTCCAAACATGGCCACAGATCACTG 5340
Db      5400 TTCAGGGGGCTAGAAATCTGTCTATGAGGCCCTTCCAAACATGGCCACAGATCACTG 5459
OY      5341 GAATGATGAGTACAGCTGTGTGCTCTGTGTGTGAAGAGCTTTTCATTCACCCCTT 5400
Db      5460 GAATGATGAGTACAGCTGTGTGCTCTGTGTGTGAAGAGCTTTTCATTCACCCCTT 5519
OY      5401 GGACACAGTGTCCACCCCAATTTGTGTTGTGCACACAGATCCCTGGACAGAGACATGGC 5460
Db      5520 GGACACAGTGTCCACCCCAATTTGTGTTGTGCACACAGATCCCTGGACAGAGACATGGC 5579
OY      5461 TTCATGATGATTTGGGAGATGTGTGAGGACCTGTGTGTGACCCGAGAGTGGTGTGGAC 5520
Db      5580 TTCATGATGATTTGGGAGATGTGTGAGGACCTGTGTGTGACCCGAGAGTGGTGTGGAC 5639
OY      5521 AGTGTAGACTCTACCAAGTGTGTGAGGACCTGTGTGTGACCCGAGAGTGGTGTGGAC 5580
Db      5640 AGTGTAGACTCTACCAAGTGTGTGAGGACCTGTGTGTGACCCGAGAGTGGTGTGGAC 5699
OY      5581 AGCCACTAC 5589
Db      5700 AGCCACTAC 5708

RESULT 10
AAV46461
ID      AAV46461 standard; cDNA; 5711 BP.
XX      AAV46461;
AC      AAV46461;
XX      18-NOV-1998 (first entry)
DE      Human BRCA1 omi2 polymorphism #4 cDNA.
XX      BRCA1; omi2; human; breast and ovarian cancer predisposing gene;
KM      polymorphism; susceptibility; anti-oncogene; tumour suppressor;
XX      chromosome 17q; ss.
XX      Homo sapiens.
XX      OS
FH      Key      Location/Qualifiers
FT      CDS      120..5711
FT      /*tag= a
FT      /*product= "BRCA1 omi2 protein"
FT      variation 3232
FT      /*tag= b
FT      /*note= "This polymorphic variation can be an A or G
XX      US5750400-A.
XX      12-MAY-1998.
XX

```

```

PF      12-FEB-1997; 97US-0798691.
XX
PR      12-FEB-1996; 96US-0598591.
PR      12-FEB-1997; 97US-0798691.
XX
PA      (ONCO-) ONCORMED INC.
XX
PI      Allen AC, Alvares CP, Critz BS, Murphy PD, Olson SJ;
PI      Schelter DB, Zeng B;
XX
DR      WPI: 1998-296774/26.
XX
PT      BRCA1 omi gene coding sequences - useful for distinguishing between
PT      polymorphisms and mutation(s) in the screening for disposition to
PT      breast or ovarian cancer
XX
PS      Claim 2e: Page -: 54pp; English.
XX
CC      This sequence encodes a human BRCA1 (breast and ovarian cancer
CC      predisposing gene) omi2 gene in which a polymorphic variation occurs at
CC      nucleotide 3232. This sequence and other polymorphic variations of this
CC      sequence are useful for the identification of an individual who may or
CC      may not have an increased susceptibility to breast or ovarian cancer.
CC      The sequences used identify gene changes which are due to polymorphisms,
CC      rather than cancer causing mutations. BRCA1 is an anti-oncogene (tumour
CC      suppressor) which is involved in genetic inheritance of cancers,
CC      especially breast and ovarian cancer. It is found at human chromosome
CC      17q which is known to be linked to cancer susceptibility, especially
CC      breast cancer. Cells containing a mutation in this gene lose the
CC      wild-type function of BRCA1 and are more susceptible to cancers.
CC      NOTE: This sequence does not appear in the specification but has been
CC      created from the wild type BRCA1 omi2 gene represented in AAV46449.
XX
SQ      Sequence 5711 BP; 1955 A; 1098 C; 1274 G; 1383 T; 1 other;

Query Match      100.0%; Score 5587; DB 19; Length 5711;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5587; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY      1 ATGGATTATCTGCTCTCCCGTGAAGAGTCAAAATGCTATATGCTATGCGAAA 60
Db      120 ATGGATTATCTGCTCTCCCGTGAAGAGTCAAAATGCTATATGCTATGCGAAA 179
OY      61 ATCTTAGAGTGTCCCATCTGTGAGTGTGATCAAGAGACCTGTCCCAAGTGTGAC 120
Db      180 ATCTTAGAGTGTCCCATCTGTGAGTGTGATCAAGAGACCTGTCCCAAGTGTGAC 239
OY      121 CACATATTTTGAATTTTGGCATGCTGAACCTCTCAACCAAGAGAAAGGGCTTCACAG 180
Db      240 CACATATTTTGAATTTTGGCATGCTGAACCTCTCAACCAAGAGAAAGGGCTTCACAG 299
OY      181 TGTCCCTTATGTAAGATGATATACCAAGAGAGCTACAAAGATACGAGATTAGT 240
Db      300 TGTCCCTTATGTAAGATGATATACCAAGAGAGCTACAAAGATACGAGATTAGT 359
OY      241 CAACCTGTTGAAGAGCATTAATGATTTGCTTTCAGCTTGACAGAGTTGGAG 300
Db      360 CAACCTGTTGAAGAGCATTAATGATTTGCTTTCAGCTTGACAGAGTTGGAG 419
OY      301 TATGCAACAGCTATATATTTTGAAGAAAGAAATTAATCTCTGTAACATCTAAAGAT 360
Db      420 TATGCAACAGCTATATATTTTGAAGAAAGAAATTAATCTCTGTAACATCTAAAGAT 479
OY      361 GAAGTTTCTATCTATCCAAATATGAGGCTACAGAAACCGTCCAAAGACTTCTACAGAGT 420
Db      480 GAAGTTTCTATCTATCCAAATATGAGGCTACAGAAACCGTCCAAAGACTTCTACAGAGT 539
OY      421 GAACCCGAAATATCTTCTGTCAGAGAAACAGAGTCTCAAGTCTCTAATCTTGA 480
Db      540 GAACCCGAAATATCTTCTGTCAGAGAAACAGAGTCTCAAGTCTCTAATCTTGA 539
OY      481 ACTGTGAGAACTGTGAGAGCAAGAGCGGATACAACTCAAGAGAGCTGTCTACATT 540

```

Db	600	ACTGTGAAACTCTGTGAGACAAGACGAGGATACACACTCAAAAGACCTCTGTACTATT	655
QY	541	GAATTGGATCTGATTTCTTCTGAGATACCGTTAATAAGCAACTATTGACGTGTGGGA	600
Db	660	GAATTGGGATCTGATTTCTTCTGAGATACCGTTAATAAGCAACTATTGACGTGTGGGA	719
QY	601	GATCAAGAATTGTTTCAAAATCACCCCTCCACGAACACAGGATGAATACGTTGGATCTC	660
Db	720	GATCAAGAATTGTTTCAAAATCACCCCTCCACGAACACAGGATGAATACGTTGGATCTC	779
QY	661	.GCAAAAAAGCGTCTTGTGAATTTTTTCTGAGACGATGTGAACAAATATCAATCATCA	720
Db	780	GCAAAAAAGCGTCTTGTGAATTTTTTCTGAGACGATGTGAACAAATATCAATCATCA	839
QY	721	CCCAATAATAATGATTGTAACACCACCTGAGAGCGTGCAGTGTGAGAGGCATCCAGAAA	780
Db	840	CCCAATAATAATGATTGTAACACCACCTGAGAGCGTGCAGTGTGAGAGGCATCCAGAAA	899
QY	781	TATCAGGGTACTTCTGTTTCAAACTTGATGTGGAGCCATGTGGCACAAATACTCATGCC	840
Db	900	TATCAGGGTACTTCTGTTTCAAACTTGATGTGGAGCCATGTGGCACAAATACTCATGCC	959
QY	841	AGCTCATATACAGCATGAGACAGCGATTATTATCACTATTAAGACAGATTAATGTAGAA	900
Db	960	AGCTCATATACAGCATGAGACAGCGATTATTATCACTATTAAGACAGATTAATGTAGAA	1019
QY	901	AAGCGTAATTCCTGTAAATMAAGCAACAGCGCTGGCTTAGCAAGAGGCCAATATACAGA	960
Db	1020	AAGCGTAATTCCTGTAAATMAAGCAACAGCGCTGGCTTAGCAAGAGGCCAATATACAGA	1079
QY	961	TGGGCTGGAAGTAAAGAAACATGTATATATAGCGGACCTCCAGCACAGAAAAGGTA	1020
Db	1080	TGGGCTGGAAGTAAAGAAACATGTATATATAGCGGACCTCCAGCACAGAAAAGGTA	1139
QY	1021	GATCTGAATGCTGATCCCCCTGTGTGAGAAAGAAATGATTAAGCAACAACTCCATATC	1080
Db	1140	GATCTGAATGCTGATCCCCCTGTGTGAGAAAGAAATGATTAAGCAACAACTCCATATC	1199
QY	1081	TCAGAGAACTCTAGAGATACTGAAGATGTTCTTGGATACACTAAATAGCAGCATTCAG	1140
Db	1200	TCAGAGAACTCTAGAGATACTGAAGATGTTCTTGGATACACTAAATAGCAGCATTCAG	1259
QY	1141	AAAGTTAATGATGCTGTTTCCAGAAAGTATTAACCTGTTAGTTCTGATGACTCAATGAT	1200
Db	1260	AAAGTTAATGATGCTGTTTCCAGAAAGTATTAACCTGTTAGTTCTGATGACTCAATGAT	1319
QY	1201	GGGGAGTCTGATCAAAATGCCAAAGTAGCTATGATTGGACGTTCTTAATAGAGGTAGAT	1260
Db	1320	GGGGAGTCTGATCAAAATGCCAAAGTAGCTATGATTGGACGTTCTTAATAGAGGTAGAT	1379
QY	1261	GAATATTTCTGTTCTTCAGAGAAAATAGACTTACTGGCCACAGTATCTCATGAGGCTTTA	1320
Db	1380	GAATATTTCTGTTCTTCAGAGAAAATAGACTTACTGGCCACAGTATCTCATGAGGCTTTA	1439
QY	1321	ATATGTAAAGTGAAGAGTTCACTCCAATCACTAGAGAGTAATATTGAACACAAATA	1380
Db	1440	ATATGTAAAGTGAAGAGTTCACTCCAATCACTAGAGAGTAATATTATTGAAGACAAATA	1499
QY	1381	TTTGGGAAAACCTTGGGAAGAGCAAGGCTCCCAACTTAACCCATGTATACGTGAANAAT	1440
Db	1500	TTTGGGAAAACCTTGGGAAGAGCAAGGCTCCCAACTTAACCCATGTATACGTGAANAAT	1559
QY	1441	CTAATTATAGAGCATTTGTTACTGAGCCACAGATTAATATACAAAGAGCGTCCCTCACAAAT	1500
Db	1560	CTAATTATAGAGCATTTGTTACTGAGCCACAGATTAATATACAAAGAGCGTCCCTCACAAAT	1619
QY	1501	AAATTTAAAGCGTTAAAGAGAGACCTACATCAGGCTTCATCCTGAGGATTTTATCAAGAAA	1560
Db	1620	AAATTTAAAGCGTTAAAGAGAGACCTACATCAGGCTTCATCCTGAGGATTTTATCAAGAAA	1679
QY	1561	GCAGATTTGGCAGTTCACAAAGACCTCTGAAATGTAAATCAGGGAACCTAACCAACGGAG	1620
Db	1680	GCAGATTTGGCAGTTCACAAAGACCTCTGAAATGTAAATCAGGGAACCTAACCAACGGAG	1739

OY	162	CAGAAATGTCAGATGATATTTACCAATAGTGGTCATGAGAAATPAAACCAAGCTAT	1680
Db	1740	CAGATGGTCACTGATGATATTTACTAATAGTGGTCATGAGAAATPAAACCAAGCTAT	1799
OY	1681	TCTATTGAGATGAGAAAAATCCTAACCCATAGAAATCACTGGAAAAAGAAATCTCTTC	1740
Db	1800	TCTATTGAGATGAGAAAAATCCTAACCCATAGAAATCACTGGAAAAAGAAATCTCTTC	1859
OY	1741	AAACGGAAGCTGACCTATATACGACGATATAGCAATATGCACTGCAATTAATATTC	1800
Db	1860	AAACGGAAGCTGACCTATATACGACGATATAGCAATATGCACTGCAATTAATATTC	1919
OY	1801	CACAAATTCAAAGCACCCTAAAGAAAGATAGCGTGGAGGAGAGCTTCTTACAGGCATATT	1860
Db	1920	CACAAATTCAAAGCACCCTAAAGAAAGATAGCGTGGAGGAGAGCTTCTTACAGGCATATT	1979
OY	1861	CATCGCGTTGAACTAGTATGTCAGTAATCTTAAGCCCACTTAATTTGATCTGAATTCGA	1920
Db	1980	CATCGCGTTGAACTAGTATGTCAGTAATCTTAAGCCCACTTAATTTGATCTGAATTCGA	2039
OY	1921	ATTGATAGTTGTTCTTACAGTGAAGAGATPAAAGAAAAAAGTACACCAATGGCAGTC	1980
Db	2040	ATTGATAGTTGTTCTTACAGTGAAGAGATPAAAGAAAAAAGTACACCAATGGCAGTC	2099
OY	1981	AGGCACAGCAGAAAACTTACACTCATGGAAGTTAAAGAACTGCAACTGGAGCCAAAG	2040
Db	2100	AGGCACAGCAGAAAACTTACACTCATGGAAGTTAAAGAACTGCAACTGGAGCCAAAG	2159
OY	2041	AGTAAACAAGCCCAATGAAACAGACAAGTAAAGAAACATGACATGATATCTTCCAGAGCTG	2100
Db	2160	AGTAAACAAGCCCAATGAAACAGACAAGTAAAGAAACATGACATGATATCTTCCAGAGCTG	2219
OY	2101	AAGTTAAACAATGACACCTGGTCTTTTACTAAGTGTTCAAATACCGAGTAATTAAGAA	2160
Db	2220	AAGTTAAACAATGACACCTGGTCTTTTACTAAGTGTTCAAATACCGAGTAATTAAGAA	2279
OY	2161	TTTGTCAATCTTACGCTTCCAGAGAAAGAAAGAGAAACTGAAACAGTTAAAGTG	2220
Db	2280	TTTGTCAATCTTACGCTTCCAGAGAAAGAAAGAGAAACTGAAACAGTTAAAGTG	2339
OY	2221	TCTAATATGCTGAAGACCCCAAGATATCATGTTAAGTGGAGAAAGGTTTTGGAACT	2280
Db	2340	TCTAATATGCTGAAGACCCCAAGATATCATGTTAAGTGGAGAAAGGTTTTGGAACT	2399
OY	2281	GAAAGATCTGTAGAGATGACAGTATTTCAATTTGTAACCTGGTACTGATTTATGGCACTG	2340
Db	2400	GAAAGATCTGTAGAGATGACAGTATTTCAATTTGTAACCTGGTACTGATTTATGGCACTG	2459
OY	2341	GAAAGTATCTGTTACTGGAAGTTACACCTTAGGGAAGGCAAAAACGAACCAATTA	2400
Db	2460	GAAAGTATCTGTTACTGGAAGTTACACCTTAGGGAAGGCAAAAACGAACCAATTA	2519
OY	2401	TGTGTGAGTACAGTGTGAGCATTTTGAAGAACCCCAAGGAGCATTAATTCATGGTTGCCAA	2460
Db	2520	TGTGTGAGTACAGTGTGAGCATTTTGAAGAACCCCAAGGAGCATTAATTCATGGTTGCCAA	2579
OY	2461	GATATATGAATATGACACAGAGCCTTAAGTATCATTTGGGACATGAAGTTAACCACAGT	2520
Db	2580	GATATATGAATATGACACAGAGCCTTAAGTATCATTTGGGACATGAAGTTAACCACAGT	2639
OY	2521	CGGGAACAAGCAATPACAAAATGGAAGAAATGGAATGTAATGCTCATATTTTGCAGATACA	2580
Db	2640	CGGGAACAAGCAATPACAAAATGGAAGAAATGGAATGTAATGCTCATATTTTGCAGATACA	2699
OY	2581	TTCAAGGTTTCAAAGGCCAGTCAATTTGCTCTGTTTCAAATCCAGAAATGCGAAGAG	2640
Db	2700	TTCAAGGTTTCAAAGGCCAGTCAATTTGCTCTGTTTCAAATCCAGAAATGCGAAGAG	2759
OY	2641	GAATGTGCAACATTTCTCTGCCACATCTGGGTCTTAAAGAAACAAGTCCAAAGTCACT	2700
Db	2760	GAATGTGCAACATTTCTCTGCCACATCTGGGTCTTAAAGAAACAAGTCCAAAGTCACT	2819

OY 2701 TTGGAATGTGACAAAAGAGAAATCAAGAAAGATAGTCTAATATCAAGCCCTGTA 2760  
 DB 2820 TTGGAATGTGACAAAAGAGAAATCAAGAAAGATAGTCTAATATCAAGCCCTGTA 2879  
 OY 2761 CAGACATTAATATCACTGCAAGCTTTCCTGCTGTTGTCAGAAAGATTAAGCAATTGAT 2820  
 DB 2880 CAGACATTAATATCACTGCAAGCTTTCCTGCTGTTGTCAGAAAGATTAAGCAATTGAT 2939  
 OY 2821 AATGCCAAATGTATGATCAAGAGAGCTAGGTTTGTCTATCATCTGCTCAAGGCG 2880  
 DB 2940 AATGCCAAATGTATGATCAAGAGAGCTAGGTTTGTCTATCATCTGCTCAAGGCG 2999  
 OY 2881 AAGGAACCTGACATTAATCAAAATTAACATGAGACTTTTACAAAACCATATCGTATA 2940  
 DB 3000 AAGGAACCTGACATTAATCAAAATTAACATGAGACTTTTACAAAACCATATCGTATA 3059  
 OY 2941 CCACACCTTTTCCCATCAAGTCATTTGTTAAACTAATGTAAAGAAAATCTGCTAGAG 3000  
 DB 3060 CCACACCTTTTCCCATCAAGTCATTTGTTAAACTAATGTAAAGAAAATCTGCTAGAG 3119  
 OY 3001 GAAACCTTTGAGGAACATTCATGTCTACCTGAAAGAGAAATGGGAAATGAGAACATTCGA 3060  
 DB 3120 GAAACCTTTGAGGAACATTCATGTCTACCTGAAAGAGAAATGGGAAATGAGAACATTCGA 3179  
 OY 3061 AGTACAGTGAACAATTTAGCCCTAATACATTAGAGAAAAATGTTTAAAGAACCCAGC 3120  
 DB 3180 AGTACAGTGAACAATTTAGCCCTAATACATTAGAGAAAAATGTTTAAAGAACCCAGC 3239  
 OY 3121 TCAGACAAATTTAATGACAGTGTCTCAGTACTAATGAAATGGGCTCCAGTATTATGAA 3180  
 DB 3240 TCAGACAAATTTAATGACAGTGTCTCAGTACTAATGAAATGGGCTCCAGTATTATGAA 3299  
 OY 3181 ATAGGTTCCAGTATGAAAAATCATTCAGAGAACTAGTGAAGAACAGAGGCGCAAAATTG 3240  
 DB 3300 ATAGGTTCCAGTATGAAAAATCATTCAGAGAACTAGTGAAGAACAGAGGCGCAAAATTG 3359  
 OY 3241 AATGCTATGCTATGATTAAGGGGTTTTCGCAACCTGAGGCTCTAATAACAAAGCTTCTCGTA 3300  
 DB 3360 AATGCTATGCTATGATTAAGGGGTTTTCGCAACCTGAGGCTCTAATAACAAAGCTTCTCGTA 3419  
 OY 3301 AGTAATTTGATAGATCTCTGAAATTAATAAAGCAAGATATGAAGACTAGTTCAGACTGTT 3360  
 DB 3420 AGTAATTTGATAGATCTCTGAAATTAATAAAGCAAGATATGAAGACTAGTTCAGACTGTT 3479  
 OY 3361 AATACAGATTTCTCTCATATCTGATTTGATAGTACTAGTAAGAGCTATGGGAAAGT 3420  
 DB 3480 AATACAGATTTCTCTCATATCTGATTTGATAGTACTAGTAAGAGAGCTATGGGAAAGT 3539  
 OY 3421 CATGCACTCAGGTTTGTCTGAGACACCTGATGACCTGTTAGATGATGGTGAATAAG 3480  
 DB 3540 CATGCACTCAGGTTTGTCTGAGACACCTGATGACCTGTTAGATGATGGTGAATAAG 3599  
 OY 3481 GAGATATGATAGTCTCTGAAATTAATAAAGCAAGATATGAAGACTAGTTCAGACTGTT 3540  
 DB 3600 GAGATATGATAGTCTCTGAAATTAATAAAGCAAGATATGAAGACTAGTTCAGACTGTT 3659  
 OY 3541 GTCCAGAAAAGAGAGCTTAGCAGAGAGTCTAGGCTTTTACCCATACATTTGGGCTCAG 3600  
 DB 3660 GTCCAGAAAAGAGAGCTTAGCAGAGAGTCTAGGCTTTTACCCATACATTTGGGCTCAG 3719  
 OY 3601 GGTACCGAAGAGAGGCGCAAGAAATTAAGTCTCTGAGAGAGAACTATCTAGTGAAGAT 3660  
 DB 3720 GGTACCGAAGAGAGGCGCGCAAGAAATTAAGTCTCTGAGAGAGAACTATCTAGTGAAGAT 3779  
 OY 3661 GAAGACCTTCCCTGCTTCAACACTGTGTTATTTGGTAAAGTAAACAATATACCTTCTCAG 3720  
 DB 3780 GAAGACCTTCCCTGCTTCAACACTGTGTTATTTGGTAAAGTAAACAATATACCTTCTCAG 3839  
 OY 3721 TCTACTAGCATAGCACCCTGCTAGCAGAGTCTGCTAAGAAACAGAGAGGAATTTA 3780  
 DB 3840 TCTACTAGCATAGCACCCTGCTAGCAGAGTCTGCTAAGAAACAGAGAGGAATTTA 3899  
 OY 3781 TTATCATTTGAAGAAATGCTTAAATGACTGACGTACACAGGTAAATTTGGCGAAAGGACATCT 3840

DB 3900 TTATCATTTGAAGATATACCTTAATAGACTGCAGTACACAGGAATATTTGGCAAGGACATCT 3959  
 OY 3841 CAGGACATCACTTATAGTGAGGAAACAAATGTTCTGCTAGCTGTTTCTTCACAGTGC 3900  
 DB 3960 CAGGACATCACTTATAGTGAGGAAACAAATGTTCTGCTAGCTGTTTCTTCACAGTGC 4019  
 OY 3901 AGTAATTTGGAACACTGACTGCAAAATACAAACCCAGGATCTTCTGATTTGTTCT 3960  
 DB 4020 AGTAATTTGGAACACTGACTGCAAAATACAAACCCAGGATCTTCTGATTTGTTCT 4079  
 OY 3961 TCCAAACAAATGAGGCTACGCTGGAAGCCAGGAGGAGTGTCTGAGTCAAGAAATTTG 4020  
 DB 4080 TCCAAACAAATGAGGCTACGCTGGAAGCCAGGAGGAGTGTCTGAGTCAAGAAATTTG 4139  
 OY 4021 GTTTCAGATGATGAGAAAGAGAGAGCGGCTTGAAGAAATTAATCAAGAGAGCAAGC 4080  
 DB 4140 GTTTCAGATGATGAGAAAGAGAGAGCGGCTTGAAGAAATTAATCAAGAGAGCAAGC 4199  
 OY 4081 ATGGATTTCAAACTTAGGTGAAGAGCATCTGGGTGTGAGTGAACCAACGCTCTCGAA 4140  
 DB 4200 ATGGATTTCAAACTTAGGTGAAGAGCATCTGGGTGTGAGTGAACCAACGCTCTCGAA 4259  
 OY 4141 GACTGCTCAGGGCTATCTCTCAGAGTACATTTTAACTCAGCAGAGAGGATACCATG 4200  
 DB 4260 GACTGCTCAGGGCTATCTCTCAGAGTACATTTTAACTCAGCAGAGAGGATACCATG 4319  
 OY 4201 CAACATTAATCTGATTAAGGCTCCAGCAGAAATGCTGTAAGTGTGTTAGAACAG 4260  
 DB 4320 CAACATTAATCTGATTAAGGCTCCAGCAGAAATGCTGTAAGTGTGTTAGAACAG 4379  
 OY 4261 CATGGGAGCGGCTTCTACACGCTCCCTTCATATAGTACCTCTGCGCTTGAG 4320  
 DB 4380 CATGGGAGCGGCTTCTACACGCTCCCTTCATATAGTACCTCTGCGCTTGAG 4439  
 OY 4321 GACCTCGGAATCCAGAACCAAAACACATCAGAAAAAGCAATATTAATCTTCAGAAAAAGT 4380  
 DB 4440 GACCTCGGAATCCAGAACCAAAACACATCAGAAAAAGCAATATTAATCTTCAGAAAAAGT 4499  
 OY 4381 AGTGAATACCTTATTAAGCCAGAAATCCAGAAAGGCTTCTGCTCAAGAGTTGAGGTGCT 4440  
 DB 4500 AGTGAATACCTTATTAAGCCAGAAATCCAGAAAGGCTTCTGCTCAAGAGTTGAGGTGCT 4559  
 OY 4441 GCAGATAGTCTCTCCAGTAAATAAAGAACCGAGAGTGAAGAGTATCCCTCTCAAA 4500  
 DB 4560 GCAGATAGTCTCTCCAGTAAATAAAGAACCGAGAGTGAAGAGTATCCCTCTCAAA 4619  
 OY 4501 TGCCCATCATTAATGATAGTGTGTGATGATGACAGTGTGCTGGAGTCTTCAGAAATAGA 4560  
 DB 4620 TGCCCATCATTAATGATAGTGTGTGATGATGACAGTGTGCTGGAGTCTTCAGAAATAGA 4679  
 OY 4561 AACTACCATCTCAAGAGAGCTCATTAAAGGTTGTGTGATGATGAGAGCAACAGCTGGAA 4620  
 DB 4680 AACTACCATCTCAAGAGAGCTCATTAAAGGTTGTGTGATGATGAGAGCAACAGCTGGAA 4739  
 OY 4621 GAGTCTGGGCGCAACGATTTGAGGGAACATCTTACTTGGCAAGGGAAGTCTAGAGGGA 4680  
 DB 4740 GAGTCTGGGCGCAACGATTTGAGGGAACATCTTACTTGGCAAGGGAAGTCTAGAGGGA 4799  
 OY 4681 ACCCCTTACCTGGAATCTGGAATCAACCTCTTCTGATGACCCCTGAATCTGATCTCT 4740  
 DB 4800 ACCCCTTACCTGGAATCTGGAATCAACCTCTTCTGATGACCCCTGAATCTGATCTCT 4859  
 OY 4741 GAAGACAGAGCCCGCAGCTGAGTGTGTTGGCAGACATACATCTTCAACCTGCAATG 4800  
 DB 4860 GAAGACAGAGCCCGCAGCTGAGTGTGTTGGCAGACATACATCTTCAACCTGCAATG 4919  
 OY 4801 AAGGTCCCAATTTGAAGTGTGAGAAATCGCCCAAGTCCAGCTGCTCATATCTACT 4860  
 DB 4920 AAGGTCCCAATTTGAAGTGTGAGAAATCGCCCAAGTCCAGCTGCTCATATCTACT 4979  
 OY 4861 GATACTGCTGGGTATATGCAATGGAAGAAAGTGTGAGAGAGGAGAGCAGAAATTTGACA 4920

```

Db 4980 GATAGTCTGGTATATGCAATGGAAGAAAGTGTGAGCAGGAGAGCCAGATTGACA 5039
QY 4921 GGTCAACAGAAAGGGTCAACAAAGAAAGTCCATGGTGTGCTGGCCGACCAGAA 4980
Db 5040 GCTTCACAGAAAGGGTCAACAAAGAAAGTCCATGGTGTGCTGGCCGACCAGAA 5099
QY 4981 GAATTTATCTGCTGATCAAGTTTGGCAGAAAACACACATCACTTAATACTTAAT 5040
Db 5100 GAATTTATCTGCTGATCAAGTTTGGCAGAAAACACACATCACTTAATACTTAAT 5159
QY 5041 ACTGAAGAACTACTATGTTGTTATGAAGAAACAGATGCTGAGTTTGTGTGACGACA 5100
Db 5160 ACTGAAGAACTACTATGTTGTTATGAAGAAACAGATGCTGAGTTTGTGTGACGACA 5219
QY 5101 CTGAATATTTTCTAGAAATGCGGAGAGAAATGGGTAGTATTTCTGAGTGACC 5160
Db 5220 CTGAATATTTTCTAGAAATGCGGAGAGAAATGGGTAGTATTTCTGAGTGACC 5279
QY 5161 CAGTCTATTAAAGAAAGAAATGCTGAATGAGCATGATTTTGAAGTCAGAGAGATG 5220
Db 5280 CAGTCTATTAAAGAAAGAAATGCTGAATGAGCATGATTTTGAAGTCAGAGAGATG 5339
QY 5221 GTCAATGGAAGAAACCCAGAGTCCAAAGCCGAGCAAGAAATCCAGACAGAAAGATC 5280
Db 5340 GTCAATGGAAGAAACCCAGAGTCCAAAGCCGAGCAAGAAATCCAGACAGAAAGATC 5399
QY 5281 TTCAGAGGGGCTAGAAATCTGTTGCTATGGGCTTACCAACATGCGCCACAGATCAACTG 5340
Db 5400 TTCAGAGGGGCTAGAAATCTGTTGCTATGGGCTTACCAACATGCGCCACAGATCAACTG 5459
QY 5341 GAATGAGTGTACAGCTGTGTGCTGCTTGTGTGAGAGAGCTTTCAATCAATTCACCCCT 5400
Db 5460 GAATGAGTGTACAGCTGTGTGCTGCTTGTGTGAGAGAGCTTTCAATCAATTCACCCCT 5519
QY 5401 GGACAGAGTGTCCACCCATTTGCTGTGTGACAGCCAGATGCCGACAGAGCAATGCG 5460
Db 5520 GGACAGAGTGTCCACCCATTTGCTGTGTGACAGCCAGATGCCGACAGAGCAATGCG 5579
QY 5461 TTCATGCAATTTGGGAGATGTGTGAGCAGCCTGTGTGACCCGAGAGTGGGTGTGAC 5520
Db 5580 TTCATGCAATTTGGGAGATGTGTGAGCAGCCTGTGTGACCCGAGAGTGGGTGTGAC 5639
QY 5521 AGTGTAGCACTTACCAAGTCCAGAGAGTGTGACACCTTACCTGATCCCAAGATCCCCAC 5580
Db 5640 AGTGTAGCACTTACCAAGTCCAGAGAGTGTGACACCTTACCTGATCCCAAGATCCCCAC 5699
QY 5581 AGCCACTAC 5589
Db 5700 AGCCACTAC 5708

RESULT 11
AAV46462
ID AAV46462 standard; cDNA: 5711 BP.
XX AAV46462;
AC
XX
XX
XX 18-NOV-1998 (first entry)
DE Human BRCA1 omi2 polymorphism #5 cDNA.
XX
XX BRCA1; omi2; human: breast and ovarian cancer predisposing gene;
KW polymorphism; susceptibility; anti-oncogene; tumour suppressor;
KW chromosome 17q; ss.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH CDS 120..5711
FT /*tag= a
FT /*product= "BRCA1 omi2 protein"
FT variation 3667
FT /*tag= b

```

```

FT /note= "This polymorphic variation can be an A or G
FT nucleotide"
XX US5750400-A.
PN
XX
XX 12-MAY-1998.
PD
XX
XX 12-FEB-1997; 97US-0798691.
PE
XX
XX 12-FEB-1996; 96US-0598591.
PR
XX 12-FEB-1997; 97US-0798691.
PA (ONCO-) ONCORMED INC.
XX
XX Allen AC, Alvares CP, Critz BS, Murphy PD, Olson SJ;
PI Schelter DB, Zeng B;
XX
XX WPI: 1998-296774/26.
DR
XX
XX BRCA1 omi gene coding sequences - useful for distinguishing between
PT polymorphisms and mutation(s) in the screening for disposition to
PT breast or ovarian cancer
XX
XX Claim 2e: Page -: 54pp: English.
PS
XX
XX This sequence encodes a human BRCA1 (breast and ovarian cancer
CC predisposing gene) omi2 gene in which a polymorphic variation occurs at
CC nucleotide 3667. This sequence and other polymorphic variations of this
CC sequence are useful for the identification of an individual who may or
CC may not have an increased susceptibility to breast or ovarian cancer.
CC The sequences used identify gene changes which are due to polymorphisms,
CC rather than cancer causing mutations. BRCA1 is an anti-oncogene (tumour
CC suppressor) which is involved in genetic inheritance of cancers,
CC especially breast and ovarian cancer. It is found at human chromosome
CC 17q which is known to be linked to cancer susceptibility, especially
CC breast cancer. Cells containing a mutation in this gene lose the
CC wild-type function of BRCA1 and are more susceptible to cancers.
CC NOTE: This sequence does not appear in the specification but has been
CC created from the wild type BRCA1 omi2 gene represented in AAV46449.
XX
XX
SQ Sequence 5711 BP; 1955 A; 1098 C; 1274 G; 1383 T; 1 other:

Query Match 100.0%; Score 5587; DB 19; Length 5711;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5587; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGATTTATCTGCTCTTGGCTTGAAGAGTACAAATGTCATTAATGTCACAGAA 60
Db 120 ATGGATTTATCTGCTCTTGGCTTGAAGAGTACAAATGTCATTAATGTCACAGAA 179
QY 61 ATCTTAGAGTGTCCATCTGCTGTGAGATTGATCAAGAACTGTCTCCAAAGTGTGAC 120
Db 180 ATCTTAGAGTGTCCATCTGCTGTGAGATTGATCAAGAACTGTCTCCAAAGTGTGAC 239
QY 121 CACATATTTTGAATTTTGCATGCTGAACCTTCTCAACAGAGAAAGGCGCTTCACAG 180
Db 240 CACATATTTTGAATTTTGCATGCTGAACCTTCTCAACAGAGAAAGGCGCTTCACAG 299
QY 181 TGTCTTTATGTAAGATGATATTAACCAAGAGAGCTTACAGAAAGTACAGATTAGT 240
Db 300 TGTCTTTATGTAAGATGATATTAACCAAGAGAGCTTACAGAAAGTACAGATTAGT 359
QY 241 CAACTGTGTGAGAGCTATTTGAATTCATTTGCTTTTCAAGCTTGACAGGTTTGAG 300
Db 360 CAACTGTGTGAGAGCTATTTGAATTCATTTGCTTTTCAAGCTTGACAGGTTTGAG 419
QY 301 TATGCAACAGCTATTAATTTTGCAGAAAGAAATTAATCTCTGTAACATCTAAAGAT 360
Db 420 TATGCAACAGCTATTAATTTTGCAGAAAGAAATTAATCTCTGTAACATCTAAAGAT 479
QY 361 GAAGTTTCTATATCAAGATGTGGGCTACAGAAACGCTGCCAAAGACTTCTACAGAGT 420
Db 480 GAAGTTTCTATCATCCAAAGTATGGGCTACAGAAACGCTGCCAAAGACTTCTACAGAGT 539

```



QY 421 GAAACCGAAATCCTTCCTTGACAGAAACCACTGCTGACACTCTTCAACTTGA 480  
DB 540 GAAACCGAAATCCTTCCTTGACAGAAACCACTGCTGACACTCTTCAACTTGA 599  
QY 481 ACCTGAGAACTCTGAGAGAACAGCGGATACAACTTCAAAAGAGCTGTCTACAT 540  
DB 600 ACCTGAGAACTCTGAGAGAACAGCGGATACAACTTCAAAAGAGCTGTCTACAT 659  
QY 541 GAATGGAGATCTGATCTCTGAGAGATACCGTTAATAGCACTTATTCAGTGGGA 600  
DB 660 GAATGGAGATCTGATCTCTGAGAGATACCGTTAATAGCACTTATTCAGTGGGA 719  
QY 601 GATCAAGAAATTTGACAAATCACCCCTCAAGAACAGGATGAAATCAGTTGGATTC 660  
DB 720 GATCAAGAAATTTGACAAATCACCCCTCAAGAACAGGATGAAATCAGTTGGATTC 779  
QY 661 GCAAAAAAGGCTGCTGTGAAATTTCTGAGACGGATGTAACTAACTGACATCAA 720  
DB 780 GCAAAAAAGGCTGCTGTGAAATTTCTGAGACGGATGTAACTAACTGACATCAA 839  
QY 721 CCCAGTAATATGATTTGAACACCACTGAGACGCTGACCTGAGAGGCTCCAGAAAG 780  
DB 840 CCCAGTAATATGATTTGAACACCACTGAGACGCTGAGAGGCTCCAGAAAG 899  
QY 781 TATCAGGATGATCTGTTTCAAACTTGCAATGTGAGCCATGTGGCAAAATCTCATGCC 840  
DB 900 TATCAGGATGATCTGTTTCAAACTTGCAATGTGAGCCATGTGGCAAAATCTCATGCC 959  
QY 841 AGCTCATATCAGCATGAGAACAGCAGTTTATTAATCTCTAAAGACAGATGAATGAGAA 900  
DB 960 AGCTCATATCAGCATGAGAACAGCAGTTTATTAATCTCTAAAGACAGATGAATGAGAA 1019  
QY 901 AAGGCTGAATTCGTATATAAGCAAAACGCTGGCTTAGCAAGAGCCCAACTAACAGA 960  
DB 1020 AAGGCTGAATTCGTATATAAGCAAAACGCTGGCTTAGCAAGAGCCCAACTAACAGA 1079  
QY 961 TGGGCTGGAAGTAAAGAAACATGTAATGATAGGCGGACCTCCAGCAGAGAAAAAGGTA 1020  
DB 1080 TGGGCTGGAAGTAAAGAAACATGTAATGATAGGCGGACCTCCAGCAGAGAAAAAGGTA 1139  
QY 1021 GATCTGAATCTGATCCCTGCTGTGAGAGAAAAAGATGAAATAGCAGAACTGCCATGC 1080  
DB 1140 GATCTGAATCTGATCCCTGCTGTGAGAGAAAAAGATGAAATAGCAGAACTGCCATGC 1199  
QY 1081 TCAGAGATCTAGAGATGATGATGCTTGGATACACTAAATAGCAGCATTCAG 1140  
DB 1200 TCAGAGATCTAGAGATGATGATGCTTGGATACACTAAATAGCAGCATTCAG 1259  
QY 1141 AAGATTAATGATGCTTTTCCAGAGTGAATGATGATGATGATGATGATGATGATGAT 1200  
DB 1260 AAGATTAATGATGCTTTTCCAGAGTGAATGATGATGATGATGATGATGATGATGAT 1319  
QY 1201 GGGGAGCTGTAATCAAAATGCCAAAGTACGTGATGATGATGATGATGATGATGATGAT 1260  
DB 1320 GGGGAGCTGTAATCAAAATGCCAAAGTACGTGATGATGATGATGATGATGATGATGAT 1379  
QY 1261 GAATATTCGCTCTTCCAGAGAAATAGATGATGATGATGATGATGATGATGATGATGAT 1320  
DB 1380 GAATATTCGCTCTTCCAGAGAAATAGATGATGATGATGATGATGATGATGATGATGAT 1439  
QY 1321 ATATGTAAGTGAAGAGTTCACCTCAAAATCAGTAGAGATGATGATGATGATGATGATGAT 1380  
DB 1440 ATATGTAAGTGAAGAGTTCACCTCAAAATCAGTAGAGATGATGATGATGATGATGATGAT 1499  
QY 1381 TTTGGGAAAACTATGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440  
DB 1500 TTTGGGAAAACTATGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1559  
QY 1441 CTAATTAATGAGCATTTGTTACTGAGCCACAGATATATACAGAGGCTCCCTCCACAAAT 1500  
DB 1560 CTAATTAATGAGCATTTGTTACTGAGCCACAGATATATACAGAGGCTCCCTCCACAAAT 1619

QY 1501 AAATTAAGCGTAAAGAGACCTACATCAGGCTTCATCTGAGATTTTATCAAGAAA 1560  
DB 1620 AAATTAAGCGTAAAGAGACCTACATCAGGCTTCATCTGAGATTTTATCAAGAAA 1679  
QY 1561 GCAGATTTGGCAGTTCAAAGACTCCCTGGAATGATAATCAAGGAACATAACCAAGGAG 1620  
DB 1680 GCAGATTTGGCAGTTCAAAGACTCCCTGGAATGATAATCAAGGAACATAACCAAGGAG 1739  
QY 1621 CAGAAATGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680  
DB 1740 CAGAAATGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1799  
QY 1681 TCTATTCAGATGAGAAAAATCCTAACCCATAGATCACTCGAAAAAGATCTCTTTC 1740  
DB 1800 TCTATTCAGATGAGAAAAATCCTAACCCATAGATCACTCGAAAAAGATCTCTTTC 1859  
QY 1741 AAAACGAAACCTGACCTATAGCAGAGATGATGATGATGATGATGATGATGATGATGAT 1800  
DB 1860 AAAACGAAACCTGACCTATAGCAGAGATGATGATGATGATGATGATGATGATGATGAT 1919  
QY 1801 CACAAATCAAAAGCACCTTAAAGAAATAGGCTGAGAGAGTCTTCTACAGCATATT 1860  
DB 1920 CACAAATCAAAAGCACCTTAAAGAAATAGGCTGAGAGAGTCTTCTACAGCATATT 1979  
QY 1861 CATGCGCTGAGCTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920  
DB 1980 CATGCGCTGAGCTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2039  
QY 1921 ATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980  
DB 2040 ATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2099  
QY 1981 AGGACAGCAGAAACCTACACATCATGGAAGTAAAGAACTCCTGACCTGAGCCAGAAAG 2040  
DB 2100 AGGACAGCAGAAACCTACACATCATGGAAGTAAAGAACTCCTGACCTGAGCCAGAAAG 2159  
QY 2041 AGTAACAGCCCAATGAAACAGACAGTAAAGAAAGATGACAGTACTTCTCCAGAGCTG 2100  
DB 2160 AGTAACAGCCCAATGAAACAGACAGTAAAGAAAGATGACAGTACTTCTCCAGAGCTG 2219  
QY 2101 AAGTAAACAATGACCTGCTTTTCTTACTAATGATGATGATGATGATGATGATGATGATGAT 2160  
DB 2220 AAGTAAACAATGACCTGCTTTTCTTACTAATGATGATGATGATGATGATGATGATGATGAT 2279  
QY 2161 TTTGTCATCTGACCTTCCAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 2220  
DB 2280 TTTGTCATCTGACCTTCCAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 2339  
QY 2221 TCTAATTAATGCTGAAAGCCCAAGATCTCATGTTAAGTGGAGAAAGGTTTGGCAAACT 2280  
DB 2340 TCTAATTAATGCTGAAAGCCCAAGATCTCATGTTAAGTGGAGAAAGGTTTGGCAAACT 2399  
QY 2281 GAAAGATCTGTAGAGATGACAGTATTTCAATGCTGCTGATGATGATGATGATGATGATGAT 2340  
DB 2400 GAAAGATCTGTAGAGATGACAGTATTTCAATGCTGCTGATGATGATGATGATGATGATGAT 2459  
QY 2341 GAAAGATCTGTAGAGATGACAGTATTTCAATGCTGCTGATGATGATGATGATGATGATGAT 2400  
DB 2460 GAAAGATCTGTAGAGATGACAGTATTTCAATGCTGCTGATGATGATGATGATGATGATGAT 2519  
QY 2401 TGTGTAGTCACTGCTCAGCATTTGAAACCCCAAGGAGTAAATTCATGCTTTCACAA 2460  
DB 2520 TGTGTAGTCACTGCTCAGCATTTGAAACCCCAAGGAGTAAATTCATGCTTTCACAA 2579  
QY 2461 GATTAATGAATGACACAGAGGCTTTAATGATCAATTTGGGACATGATGATGATGATGATGAT 2520  
DB 2580 GATTAATGAATGACACAGAGGCTTTAATGATCAATTTGGGACATGATGATGATGATGATGAT 2639  
QY 2521 CGGGAACAAGCTATGAAATGGAAGAAAGTGAATGATGATGATGATGATGATGATGATGAT 2580  
DB 2640 CGGGAACAAGCTATGAAATGGAAGAAAGTGAATGATGATGATGATGATGATGATGATGAT 2699  
QY 2581 TTCAGGTTTCAAGCGCCAGTCAATTTGCTGTTTCAAAATCCAGAAATGACGAAGAG 2640



```
|||||
Db 2700 TTCAGGTTTCAAAAGGCCAGTCTATTGCTGTTTCAATCCAGAAATGCAAGAG 2759
QY 2641 GAATGTCGAACATCTCTCCGCCACTCTGGGTCCTTAAAGAAACAAGTCCAAAAGTACT 2700
Db 2760 GAATGTCGAACATCTCTCCGCCACTCTGGGTCCTTAAAGAAACAAGTCCAAAAGTACT 2819
QY 2701 TTGTAATGTAACAAAAGAGAGAAATCAAGAAAGATGAGTCTATATCAAGCCTGTA 2760
Db 2820 TTGTAATGTAACAAAAGAGAGAAATCAAGAAAGATGAGTCTATATCAAGCCTGTA 2879
QY 2761 CAGACAGTTAATATCACTGAGGCTTCTGTTGGTTCAGAAAGATAAGCCAGTTGAT 2820
Db 2880 CAGACAGTTAATATCACTGAGGCTTCTGTTGGTTCAGAAAGATAAGCCAGTTGAT 2939
QY 2821 AATGCCAATATGATATCAAGAGAGGCTAGAGTTTGTCTATCATCTCAAGTCAAGGC 2880
Db 2940 AATGCCAATATGATATCAAGAGAGGCTAGAGTTTGTCTATCATCTCAAGTCAAGGC 2999
QY 2881 AACGAACATGAGACTCTTACTCCAAATTAACATGAGCTTTTACAAAACCATATCGTATA 2940
Db 3000 AACGAACATGAGACTCTTACTCCAAATTAACATGAGCTTTTACAAAACCATATCGTATA 3059
QY 2941 CCACCACTTTTCCCATCAAGTCAATTTGTTAAACTAAATGTAAGAAAATCTGCTAGAG 3000
Db 3060 CCACCACTTTTCCCATCAAGTCAATTTGTTAAACTAAATGTAAGAAAATCTGCTAGAG 3119
QY 3001 GAAAACCTTGAGAAACATTCATGTCACCTGAAAGAAATGGAAATGGAACATTTCCA 3060
Db 3120 GAAAACCTTGAGAAACATTCATGTCACCTGAAAGAAATGGAAATGGAACATTTCCA 3179
QY 3061 AGTACAGTGAAGACATTTAGCCGTAATAACATTTAGAGAAAATGTTTTAAAGAGCCAGC 3120
Db 3180 AGTACAGTGAAGACATTTAGCCGTAATAACATTTAGAGAAAATGTTTTAAAGAGCCAGC 3239
QY 3121 TCAACGATATTAATGAACTAGTTCACATCTAATGAAGTGGGCTCCAGTATTAATGAA 3180
Db 3240 TCAACGATATTAATGAACTAGTTCACATCTAATGAAGTGGGCTCCAGTATTAATGAA 3299
QY 3181 ATAGGTTCCAGTGAAGAAACATTCAGAGACATGTAAGAACAGAGGGCCCAAAATG 3240
Db 3300 ATAGGTTCCAGTGAAGAAACATTCAGAGACATGTAAGAACAGAGGGCCCAAAATG 3359
QY 3241 AATGCTATGCTTATAGTTAGGGTTTTGCAACCTGAGTCTATTAACAAAGTCTTCTGGA 3300
Db 3360 AATGCTATGCTTATAGTTAGGGTTTTGCAACCTGAGTCTATTAACAAAGTCTTCTGGA 3419
QY 3301 AGTAATTTGTAACATCTGAAATTAAGAAAGCAAGAAATGAAAGAGTACTGAGACTGTT 3360
Db 3420 AGTAATTTGTAACATCTGAAATTAAGAAAGCAAGAAATGAAAGAGTACTGAGACTGTT 3479
QY 3361 AATAGAGATTTCTCCATATGATGATTTAGAGAACTTGAACAGGCTATGGAAGTACT 3420
Db 3480 AATAGAGATTTCTCCATATGATGATTTAGAGAACTTGAACAGGCTATGGAAGTACT 3539
QY 3421 CATGATCTCAGAGTTGTTCTAGACACCTGATGACCTTAGATGATGTAAGTAAG 3480
Db 3540 CATGATCTCAGAGTTGTTCTAGACACCTGATGACCTTAGATGATGTAAGTAAG 3599
QY 3481 GAGATACTAGTTTGTGTAAGAAATGATTAAGGAAAGTCTGCTGTTTTAGCAAAAGC 3540
Db 3600 GAGATACTAGTTTGTGTAAGAAATGATTAAGGAAAGTCTGCTGTTTTAGCAAAAGC 3659
QY 3541 GTCCGAAAGAGAGAGCTTGAAGAGATCTAGCCCTTACACCATACATTTGGCTGAG 3600
Db 3660 GTCCGAGARAGAGAGCTTGAAGAGATCTAGCCCTTACACCATACATTTGGCTGAG 3719
QY 3601 GGTACCCGAGAGGGGCCAAGAAATTAGAGTCTCAGAGAGAACTATCTAGTGAAGAT 3660
Db 3720 GGTACCCGAGAGGGGCCAAGAAATTAGAGTCTCAGAGAGAACTATCTAGTGAAGAT 3779
QY 3661 GAAGAGCTTCCCTGCTTCCAAACATGTTATTTGTAAGTAACAATATACCTTCTCAG 3720
|||||
```

---

```
Db 3780 GAAGAGCTTCCCTGCTTCCAAACATGTTATTTGTTAAAGTAACAATATACCTTCTCAG 3839
QY 3721 TCTACTAGGATATGACACCGCTGCTACCCGAGTCTGCTGCTAACAACAGAGAGATTTA 3780
Db 3840 TCTACTAGGATATGACACCGCTGCTACCCGAGTCTGCTGCTAACAACAGAGAGATTTA 3899
QY 3781 TTATCATTTAGAAATGAGTTTAAATGACTGACATAACACAGGTAAATTTGGCAAGCATCT 3840
Db 3900 TTATCATTTAGAAATGAGTTTAAATGACTGACATAACACAGGTAAATTTGGCAAGCATCT 3959
QY 3841 CAGGAACATCACCTTAGTAGAGAAACAAATGTTCTGCTAGCTGTTTTTCTTACAGTGC 3900
Db 3960 CAGGAACATCACCTTAGTAGAGAAACAAATGTTCTGCTAGCTGTTTTTCTTACAGTGC 4019
QY 3901 AGTGATTTGGAAGACCTTGACCTGCAAAATACAAACCCAGAGATCCCTTCTTCAATGCTCT 3960
Db 4020 AGTGATTTGGAAGACCTTGACCTGCAAAATACAAACCCAGAGATCCCTTCTTCAATGCTCT 4079
QY 3961 TCCAAACAAATGAGGATGAGTCTGAAAGCCAGGAGTTGGTCTGAGTGAACAGGAATTTG 4020
Db 4080 TCCAAACAAATGAGGATGAGTCTGAAAGCCAGGAGTTGGTCTGAGTGAACAGGAATTTG 4139
QY 4021 GTTTCAGATGATGAAGAAAGAGAAAGGCTTGAAGAAATATATCAAGAAAGCAAGC 4080
Db 4140 GTTTCAGATGATGAAGAAAGAGAAAGGCTTGAAGAAATATATCAAGAAAGCAAGC 4199
QY 4081 ATGGAATTCAACTTAGTGAACAGATCTGGGCTGAGAGTGAACAAAGGCTCTGAA 4140
Db 4200 ATGGAATTCAACTTAGTGAACAGATCTGGGCTGAGAGTGAACAAAGGCTCTGAA 4259
QY 4141 GACTGCTCAGGGGCTATCTCTCAGAGTGCATTTTAAACACTCAGCAGAGGAGATACATG 4200
Db 4260 GACTGCTCAGGGGCTATCTCTCAGAGTGCATTTTAAACACTCAGCAGAGGAGATACATG 4319
QY 4201 CAACATTAACCTGATTAAGAGTCCAGCAGAGAAATGCGTGAACATGTAAGCTGTAGAACAG 4260
Db 4320 CAACATTAACCTGATTAAGAGTCCAGCAGAGAAATGCGTGAACATGTAAGCTGTAGAACAG 4379
QY 4261 CATGGAGGACAGGCTTCTTAACAGTACCTTCATCATTAAGTACTTCTGCGCCCTGAG 4320
Db 4380 CATGGAGGACAGGCTTCTTAACAGTACCTTCATCATTAAGTACTTCTGCGCCCTGAG 4439
QY 4321 GACCTGCGAAATCCAGAAACAAAGCAATCAGAAAGAGATTAATCTTCACAGAAAGT 4380
Db 4440 GACCTGCGAAATCCAGAAACAAAGCAATCAGAAAGAGATTAATCTTCACAGAAAGT 4499
QY 4381 AGTGAATACCTTATTAAGCCAGATCCAGAGGCTTTCTGCTGACAGTGTGAGTGTCT 4440
Db 4500 AGTGAATACCTTATTAAGCCAGATCCAGAGGCTTTCTGCTGACAGTGTGAGTGTCT 4559
QY 4441 GCAGATAGTTTTCACAGTAAATTAAGAACAGGAGTGAAGAGTCAATCCCTCTTAAA 4500
Db 4560 GCAGATAGTTTTCACAGTAAATTAAGAACAGGAGTGAAGAGTCAATCCCTCTTAAA 4619
QY 4501 TGCCTCATTTAGATGATAGTGTGATACATGCACATTTCTCTGGGAGTCTTCAAGATAGA 4560
Db 4620 TGCCTCATTTAGATGATAGTGTGATACATGCACATTTCTCTGGGAGTCTTCAAGATAGA 4679
QY 4561 AACTACCATCTCAAGAGAGGCTCAATTAAGTTGTTGATGTTGAGAGAGCAACAGCTGAA 4620
Db 4680 AACTACCATCTCAAGAGAGGCTCAATTAAGTTGTTGATGTTGAGAGAGCAACAGCTGAA 4739
QY 4621 GAGTCTGGGCGACAGGATTTGAGGAAACATCTTACCTGCAAGGCAAGATCTGAGGGA 4680
Db 4740 GAGTCTGGGCGACAGGATTTGAGGAAACATCTTACCTGCAAGGCAAGATCTGAGGGA 4799
QY 4681 ACCCTTACCTGGAATCTGGAATCAGCTCTTCTCTGATGACCTGATCTGATCTCTCT 4740
Db 4800 ACCCTTACCTGGAATCTGGAATCAGCTCTTCTCTGATGACCTGATCTGATCTCTCTCT 4859
QY 4741 GAAGACAGAGCCCGAGAGTCAAGTCTGTGTGGCAACATACCATCTTCAACTCTGCAATG 4800
Db 4860 GAAGACAGAGCCCGAGAGTCAAGTCTGTGTGGCAACATACCATCTTCAACTCTGCAATG 4919
|||||
```

```

QY 4801 AAAGTCCCAATTTGAAGTTCGAGATCTGCCAGATCTCCTGCTACTACT 4860
    |||
Db 4920 AAAGTCCCAATTTGAAGTTCGAGATCTGCCAGATCTCCTGCTACTACT 4879
QY 4861 GATACGCTGGGTATATGCAATGGAAGAAAGTGTGACGAGGAGAGAACCAATGACA 4920
    |||
Db 4980 GATACGCTGGGTATATGCAATGGAAGAAAGTGTGACGAGGAGAGAACCAATGACA 5039
QY 4921 GCTTCAACGAAAGGGTCAACAAAGATGTCATGGTGTGCTGACCCAGAA 4980
    |||
Db 5040 GCTTCAACGAAAGGGTCAACAAAGATGTCATGGTGTGCTGACCCAGAA 5099
QY 4981 GAATTTATGCTGCTGATACAGTTTGGCAGAAACACACATCACTTAATTAATTAAT 5040
    |||
Db 5100 GAATTTATGCTGCTGATACAGTTTGGCAGAAACACACATCACTTAATTAATTAAT 5159
QY 5041 ACTGAAGAGACTACTGATGTTGTATGAAGAACAGATGCTGAGTTGTGTGAACGACA 5100
    |||
Db 5160 ACTGAAGAGACTACTGATGTTGTATGAAGAACAGATGCTGAGTTGTGTGAACGACA 5219
QY 5101 CTGAAATATTTCTAGGAATTTGGGGAGAAATGGGTAGTTAGTATTTCTGGGTGACC 5160
    |||
Db 5220 CTGAAATATTTCTAGGAATTTGGGGAGAAATGGGTAGTTAGTATTTCTGGGTGACC 5279
QY 5161 CAGTCTATTAAAGAAAGAAATGCTGAATGACATGATTTTGAAGTCAGAGAGATGTG 5220
    |||
Db 5280 CAGTCTATTAAAGAAAGAAATGCTGAATGACATGATTTTGAAGTCAGAGAGATGTG 5339
QY 5221 GTCATATGGAAGAAACCAACCAAGTCCAAAGCAGACAGAGATCCAGAGAAAGATC 5280
    |||
Db 5340 GTCATATGGAAGAAACCAACCAAGTCCAAAGCAGACAGAGATCCAGAGAAAGATC 5399
QY 5281 TTCAGGGGGCTAGAAATCTGTTGCTATGGGCCCTTACCAACATGCCCAGATCACTG 5340
    |||
Db 5400 TTCAGGGGGCTAGAAATCTGTTGCTATGGGCCCTTACCAACATGCCCAGATCACTG 5459
QY 5341 GAATGATGTACAGCTGTGTGCTTCTGTGTGTAAGAGAGCTTTCATCATTCACCCCT 5400
    |||
Db 5460 GAATGATGTACAGCTGTGTGCTTCTGTGTGTAAGAGAGCTTTCATCATTCACCCCT 5519
QY 5401 GGCACAGGTGTCCACCAATTTGTTGTGTCACACAGATCTCTGACAGAGCAATGCC 5460
    |||
Db 5520 GGCACAGGTGTCCACCAATTTGTTGTGTCACACAGATCTCTGACAGAGCAATGCC 5579
QY 5461 TTCATGCAATTTGGCAGATGTGTGAGGACCTGTGTGACCCGAGAGTGGGTGGAC 5520
    |||
Db 5580 TTCATGCAATTTGGCAGATGTGTGAGGACCTGTGTGACCCGAGAGTGGGTGGAC 5639
QY 5521 AGTGTAGCACTTACCAAGTCCAGAGAGTGGACACCTACTGATACCCAGATCCCCAC 5580
    |||
Db 5640 AGTGTAGCACTTACCAAGTCCAGAGAGTGGACACCTACTGATACCCAGATCCCCAC 5699
QY 5581 AGCCACTAC 5589
    |||
Db 5700 AGCCACTAC 5708

```

## RESULT 12

AAT84840 standard; cDNA to mRNA, 5712 BP.

AAT84840;

21-FEB-1998 (first entry)

Human breast and ovarian cancer susceptibility gene BRCA1.

BRCA1; breast cancer; ovarian cancer; human;

tumour suppressor gene; gene therapy; receptor; ds.

Homo sapiens.

OS

XX

```

FH Key Location/Qualifiers
FT CDS 120..5711
    /Lag= a
FT /transl_except= (pos:1482..1484, aa:Asp)
FT /transl_except= (pos:1581..1583, aa:Ser)
FT /transl_except= (pos:1899..1901, aa:Glu)
FT /transl_except= (pos:1902..1904, aa:Ileu)
FT /transl_except= (pos:1905..1907, aa:Glu)
FT /transl_except= (pos:1908..1910, aa:Ileu)
FT /transl_except= (pos:1911..1913, aa:Asn)
FT /transl_except= (pos:1914..1916, aa:Ile)
FT /transl_except= (pos:1917..1919, aa:Met)
FT /transl_except= (pos:4476..4478, aa:Val)
FT /transl_except= (pos:4479..4481, aa:Ileu)
FT /transl_except= (pos:4482..4484, aa:Gln)
FT /transl_except= (pos:4689..4691, aa:Pro)

```

W09730108-A1.

21-AUG-1997.

19-FEB-1997; 97MO-US03340.

20-FEB-1996; 96US-0603753.

(UYVA-) UNIV VANDERBILT.  
(UNIM) UNIV WASHINGTON.Claire-King M, Holt JT, Jensen RA, Jetton TL, Page DL;  
Robinson-Benlion CL, Szabo CI, Thompson ME;

MPI; 1997-434733/40.

P-PSDB; AAW23286.

BRCA1 and BRCA2 tumour suppressor gene products - useful to inhibit breast and ovarian cancer cell growth and tumorigenesis, or treat gene linked hereditary or sporadic ovarian or breast cancer

Claim 13; Page 54-63; 148pp; English.

This sequence comprises a full-length BRCA1 cDNA. Genetic analysis of familial and ovarian cancer indicates that BRCA1 is a tumour suppressor gene. It encodes a 190 kDa protein (see AAW23286) that is an inhibitor of the growth and proliferation of human breast and ovarian cancer cells. DNA encoding the BRCA1 protein can thus be used in gene therapy methods for the treatment of breast and ovarian cancers. A purified BRCA1 protein can also be used to treat the BRCA1 receptor and hence to identify BRCA1 protein-mimetic agents which act on the receptor for use in breast and ovarian cancer treatment. The BRCA2 gene (see AAT84841) and BRCA2 protein (see AAW23287) have also been characterised. Methods are claimed for the isolation of BRCA1 or BRCA2 receptors, and for treating or preventing breast and (gene-linked hereditary or sporadic) ovarian cancers using BRCA1 and BRCA2 proteins and genes.

Sequence 5712 BP; 1956 A; 1099 C; 1274 G; 1383 T; 0 other;

Query Match 99.9%; Score 5585.8; DB 18; Length 5712;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 5587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 ATGGAATTATCTGCTCTGCGGTTGAAGAGTCAAAATGCTATTATGTCAGAGAA 60
    |||
Db 120 ATGGAATTATCTGCTCTGCGGTTGAAGAGTCAAAATGCTATTATGTCAGAGAA 179
QY 61 ATCTTAGAGTGTCCACTCTGCTGAGTTGATCAAGGAACCTCTCCACAAAGTGTAC 120
    |||
Db 180 ATCTTAGAGTGTCCACTCTGCTGAGTTGATCAAGGAACCTCTCTCCAAAGTGTAC 239
QY 121 CACATATTTTGCAGAAATTTTGCAGTGAACCTTCTCAACAGAGAAAGGCGCTCACAG 180
    |||

```

D 240 CACATATTTTGCNAATTTTGCATGCTGAAACCTTCTCAACCAGAAAGAGGGCTTCCAG 299  
O 181 TGCCTTTATGTAGATGATATACCAAAAGAGGCTTACAGAAAGTACAGATTAGT 240  
D 300 TGTCTTTTATGTAGAAATGATATACCAAAAGAGGCTTACAGAAAGTACAGATTAGT 359  
O 241 CAACCTTTTGAAGAGCTATTTGAAATCATTTTGTCTTTTCAGCTTGACAGAGTTTGGAG 300  
D 360 CAACCTTTTGAAGAGCTATTTGAAATCATTTTGTCTTTTCAGCTTGACAGAGTTTGGAG 419  
O 301 TATGCAACAGCTATATTTTGCNAAAAGAAATTAATCTCTCTGACATCTTAAAGAT 360  
D 420 TATGCAACAGCTATATTTTGCNAAAAGAAATTAATCTCTCTGACATCTTAAAGAT 479  
O 361 GAAGTTTCTATCATCCAAAGTATGGGTACAGAAACCGTGCCAAAGACTTCTACAGAT 420  
D 480 GAAGTTTCTATCATCCAAAGTATGGGTACAGAAACCGTGCCAAAGACTTCTACAGAT 539  
O 421 GAACCCGAAATTCCTTCTCTGAGAAACAGTCTGAGTGTCCAACTCTTAACTTTGA 480  
D 540 GAACCCGAAATTCCTTCTCTGAGAAACAGTCTGAGTGTCCAACTCTTAACTTTGA 599  
O 481 ACTGTGAGAACTCTGAGAGCAAGACGCGATACAACTTAAAGAGCTGTCTACAT 540  
D 600 ACTGTGAGAACTCTGAGAGCAAGACGCGATACAACTTAAAGAGCTGTCTACAT 659  
O 541 GAATGGAGTCTGATCTCTGAGAGTACCGTTAATAGGCACTTATGACAGTGGGA 600  
D 660 GAATGGAGTCTGATCTCTGAGAGTACCGTTAATAGGCACTTATGACAGTGGGA 719  
O 601 GATCAGAATTTGTTACAAATCACCCCTCAAGAAACAGGAGTAAATCACTTTGATCT 660  
D 720 GATCAGAATTTGTTACAAATCACCCCTCAAGAAACAGGAGTAAATCACTTTGATCT 779  
O 661 GCAAAAAAGGCTGCTGTGATTTTCTGAGACGAGTATACAAATTAATGACATCATCA 720  
D 780 GCAAAAAAGGCTGCTGTGATTTTCTGAGACGAGTATACAAATTAATGACATCATCA 839  
O 721 CCCAGTAAATGATTTGAAACACCACTGAGAAAGCTGAGAGCTGAGAGGCTCCAGAAAG 780  
D 840 CCCAGTAAATGATTTGAAACACCACTGAGAAAGCTGAGAGCTGAGAGGCTCCAGAAAG 899  
O 781 TATCAGAGGTAGTCTGCTTCAAACTTGATGTGAGCCATGTGGCAAAATTAATCTGCC 840  
D 900 TATCAGAGGTAGTCTGCTTCAAACTTGATGTGAGCCATGTGGCAAAATTAATCTGCC 959  
O 841 ACCTATTCACAGATGAGAAACAGCACTTATTAATCTACATAAGACAGAAATGATAGAA 900  
D 960 ACCTATTCACAGATGAGAAACAGCACTTATTAATCTACATAAGACAGAAATGATAGAA 1019  
O 901 AAGGCTGAATCTGTATATAAAGCAAAACAGCTGCTTACAGAGAGCCAACTAATACAGA 960  
D 1020 AAGGCTGAATCTGTATATAAAGCAAAACAGCTGCTTACAGAGAGCCAACTAATACAGA 1079  
O 961 TGGGCTGGAATGAGGAAACATGTAATGATAGGCGGAGCTCCAGCAGAGAAAAAGGTA 1020  
D 1080 TGGGCTGGAATGAGGAAACATGTAATGATAGGCGGAGCTCCAGCAGAGAAAAAGGTA 1139  
O 1021 GATCTGAATGCTGATCCCTGTGTGAGAAAAAGAAATGAAATGAGAGAACTGCCATGC 1080  
D 1140 GATCTGAATGCTGATCCCTGTGTGAGAAAAAGAAATGAAATGAGAGAACTGCCATGC 1199  
O 1081 TCAGAGAAATCTAGAGATCTAGAGATGTTCTTGTGATACACTAATATAGCAGCATTCAG 1140  
D 1200 TCAGAGAAATCTAGAGATCTAGAGATGTTCTTGTGATACACTAATATAGCAGCATTCAG 1259  
O 1141 AAAAGTTAATGAGTGTCTTCCAGAAAGTGAATCTGTAGGTTCTGATGACTACATGAT 1200  
D 1260 AAAAGTTAATGAGTGTCTTCCAGAAAGTGAATCTGTAGGTTCTGATGACTACATGAT 1319  
O 1201 GGGGAGTCTGAATCAAAATGCCAAAGTAGTGTATTTGGAGCTTCTAAATGAGTATGAT 1260  
D 1320 GGGGAGTCTGAATCAAAATGCCAAAGTAGTGTATTTGGAGCTTCTAAATGAGTATGAT 1379

O 1261 GAATATTTCTGCTTCTCAGAGAAATAGACTTACTGGCCAGATGATCCATGAGGCTTTA 1320  
D 1380 GAATATTTCTGCTTCTCAGAGAAATAGACTTACTGGCCAGATGATCCATGAGGCTTTA 1439  
O 1321 ATATGTAAGAAGTGAAGAGTTCACCTCAATCAGTATGAGAGTAATATTTGAACAAAAA 1380  
D 1440 ATATGTAAGAAGTGAAGAGTTCACCTCAATCAGTATGAGAGTAATATTTGAACAAAAA 1499  
O 1381 TTTGGGAAAACTATGGAAGAAAGCAAGCTTCCCACTTAAAGCATGTAACATAAAT 1440  
D 1500 TTTGGGAAAACTATGGAAGAAAGCAAGCTTCCCACTTAAAGCATGTAACATAAAT 1559  
O 1441 CTAAATATAGAGCATTTGTTACTGAGCCACAGATATATCAAGAGAGTCCCTCAAAAT 1500  
D 1560 CTAAATATAGAGCATTTGTTACTGAGCCACAGATATATCAAGAGAGTCCCTCAAAAT 1619  
O 1501 AAATTAAGGCTTAAAGAGACCTACATGAGGCTTCATCTGAGGATTTTATCAAGAA 1560  
D 1620 AAATTAAGGCTTAAAGAGACCTACATGAGGCTTCATCTGAGGATTTTATCAAGAA 1679  
O 1561 GCAGATTTGGCAGTTCAAAAAGACTCCTGAAATGATTAATCAGAGAACTTAACCAAGGAG 1620  
D 1680 GCAGATTTGGCAGTTCAAAAAGACTCCTGAAATGATTAATCAGAGAACTTAACCAAGGAG 1739  
O 1621 CAGATGTCACAGTATGATATTAATTAATAGTGTGATGAGATTAACCAAGAGTAT 1680  
D 1740 CAGATGTCACAGTATGATATTAATTAATAGTGTGATGAGATTAACCAAGAGTAT 1799  
O 1681 TCTATTCAGAAATGAGAAAAATCCTAACCAATAGAAATCACTGAAAAAGAAATCTCTTC 1740  
D 1800 TCTATTCAGAAATGAGAAAAATCCTAACCAATAGAAATCACTGAAAAAGAAATCTCTTC 1859  
O 1741 AAACGAAAGCTGAACTTAAAGAGTATTAAGCAATATGGAATGGAATTAATATTC 1800  
D 1860 AAACGAAAGCTGAACTTAAAGAGTATTAAGCAATATGGAATGGAATTAATATTC 1919  
O 1801 CACAATTCAAAAGCACCCTTAAAGAAATGAGTGGAGAGAACTTCTTACAGGAGATAT 1860  
D 1920 CACAATTCAAAAGCACCCTTAAAGAAATGAGTGGAGAGAACTTCTTACAGGAGATAT 1979  
O 1861 CATGCGCTTGAACCTAGTACGATGAGTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 1920  
D 1980 CATGCGCTTGAACCTAGTACGATGAGTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 2039  
O 1921 ATTGATAGTGTCTGAGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 1980  
D 2040 ATTGATAGTGTCTGAGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 2099  
O 1981 AGGACAGAGAAACCTTACATGATGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 2040  
D 2100 AGGACAGAGAAACCTTACATGATGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 2159  
O 2041 AGTAAACAGCCAAATGAAACAGCAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 2100  
D 2160 AGTAAACAGCCAAATGAAACAGCAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 2219  
O 2101 AAGTTAACAATGACACCTGCTCTTTTACTAGTGTCTTAAATACCAAGTAAAGTAAAGTAA 2160  
D 2220 AAGTTAACAATGACACCTGCTCTTTTACTAGTGTCTTAAATACCAAGTAAAGTAAAGTAA 2279  
O 2161 TTTGTCAATCTAGCTTCCAGAGAGAAAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 2220  
D 2280 TTTGTCAATCTAGCTTCCAGAGAGAAAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 2339  
O 2221 TCTAATATGCTGAGAACCCCAAGATCTCATGTTAAGTGAGAAAGGTTTCCAAACT 2280  
D 2340 TCTAATATGCTGAGAACCCCAAGATCTCATGTTAAGTGAGAAAGGTTTCCAAACT 2399  
O 2281 GAAAGATCTGTAGAGATGAGTATTTGATTTGATGTTGATGATTTATGAGTATGAT 2340  
D 2400 GAAAGATCTGTAGAGATGAGTATTTGATTTGATGTTGATGATTTATGAGTATGAT 2459

2341 GAAATATCTCGTTACTGGAAGTTAGCACTAGGAGGCAAAAACAGAACCAATAA 2400  
2460 GAAAGATATCTGTTACTGGAAGTTAGCACTAGGAGGCAAAAACAGAACCAATAA 2519  
2401 TGTGTAGTACAGTGTGACAGATTTGAAACCCCAAGGACTAATGATGTTTCCAAA 2460  
2520 TGTGTAGTACAGTGTGACAGATTTGAAACCCCAAGGACTAATGATGTTTCCAAA 2579  
2461 GATAATAGAAATGACACAGAAAGCTTTAAGTATCCATTGGACATGAACTAACACAGT 2520  
2580 GATAATAGAAATGACACAGAAAGCTTTAAGTATCCATTGGACATGAACTAACACAGT 2639  
2521 CGGGAACACAGCATAGAAATGAGAAAGTGAACCTGATGCTAGTATTTGCAAAATACA 2580  
2640 CGGGAACACAGCATAGAAATGAGAAAGTGAACCTGATGCTAGTATTTGCAAAATACA 2699  
2581 TTCAAGGTTTCAAGGGCCAGCATTTGCTCTCTTTTCAAAATCCAGAAATGCAAGAG 2640  
2700 TTCAAGGTTTCAAGGGCCAGCATTTGCTCTCTTTTCAAAATCCAGAAATGCAAGAG 2759  
2641 GAATGTGCAACATTTCTGCTCCACTGCTGGCTCTTAAAGAAACCAAGTCCAAAGTCACT 2700  
2760 GATGTGCAACATTTCTGCTCCACTGCTGGCTCTTAAAGAAACCAAGTCCAAAGTCACT 2819  
2701 TTTGAATGTGAACAAAGAGAGAAATCAAGAAAGATGAGTCTAATATCAAGCTGTA 2760  
2820 TTTGAATGTGAACAAAGAGAGAAATCAAGAAAGATGAGTCTAATATCAAGCTGTA 2879  
2761 CAGACGTTAATATCACTGCAAGCTTTCTGCTGTGTGCTCAGAAATAGCCAGTTGAT 2820  
2880 CAGACGTTAATATCACTGCAAGCTTTCTGCTGTGTGCTCAGAAATAGCCAGTTGAT 2939  
2821 AATGCCAATGTAGTATCAAGAGAGGCTAGGTTTGTCTATCATCTCAGTTCAGAGGC 2880  
2940 AATGCCAATGTAGTATCAAGAGAGGCTAGGTTTGTCTATCATCTCAGTTCAGAGGC 2999  
2881 AACGAATGTAGTATCACTGCAAAATTAACATGACCTTTTACAAAACCATATGCTATA 2940  
3000 AACGAATGTAGTATCACTGCAAAATTAACATGACCTTTTACAAAACCATATGCTATA 3059  
2941 CCACCACTTTTCCCTCAAGTATGTTTAAACATTAATGTAAGAAATGCTGCTAGAG 3000  
3060 CCACCACTTTTCCCTCAAGTATGTTTAAACATTAATGTAAGAAATGCTGCTAGAG 3119  
3001 GAAACTTTGAGGAACATTCATGTCACCTGGAAGAGAAATGGAATGAGAATTCATCCA 3060  
3120 GAAACTTTGAGGAACATTCATGTCACCTGGAAGAGAAATGGAATGAGAATTCATCCA 3179  
3061 AGTACGTAGACATTTAGCCGTAAATTAACATTTAGAGAAATGTTTTTAAAGAGCCAGC 3120  
3180 AGTACGTAGACATTTAGCCGTAAATTAACATTTAGAGAAATGTTTTTAAAGAGCCAGC 3239  
3121 TCAGCAATATTAATGAATAGGTTCCAGTACTAATGAAGTGGCCGATATTAATGAA 3180  
3240 TCAGCAATATTAATGAATAGGTTCCAGTACTAATGAAGTGGCCGATATTAATGAA 3299  
3181 ATAAGTTCCAGTATGAAACATTTCAAGCAGAACTAGTAGAAACAGAGGCCCCAAATG 3240  
3300 ATAAGTTCCAGTATGAAACATTTCAAGCAGAACTAGTAGAAACAGAGGCCCCAAATG 3359  
3241 AATGCTATGCTTAGATAGGAGGTTTTCGAACCTGAGGCTATTAACAAAGTCTCTCGA 3300  
3360 AATGCTATGCTTAGATAGGAGGTTTTCGAACCTGAGGCTATTAACAAAGTCTCTCGA 3419  
3301 AGTAATGTAGCATCTGAATATTAAGAAAGCAAGAAATGAGAACTAGTTACAGAGT 3360  
3420 AGTAATGTAGCATCTGAATATTAAGAAAGCAAGAAATGAGAACTAGTTACAGAGT 3479  
3361 AATAGCATTTCTCCATATCTGATTTAGATTAACATTAAGACAGCTATGGAAGTATG 3420  
3480 AATAGCATTTCTCCATATCTGATTTAGATTAACATTAAGACAGCTATGGAAGTATG 3539  
3421 CATGATCTCAGGTTTCTGAGACACCTGATGACCTGTTAGATGAGTGAATTAAG 3480

3540 CATGATCTCAGGTTTGTCTGAGACACCTGATAGCTGTTAGATGATGGAATTAAG 3599  
3481 GAAGATATCTAGTTTGTCTGAAATGACATTAAGAAAGTCTGCTGTTTACCAAAAGC 3540  
3600 GAAGATATCTAGTTTGTCTGAAATGACATTAAGAAAGTCTGCTGTTTACCAAAAGC 3659  
3541 GTCCAGAAAGAGAGCTTACAGAGTCTAGGCTTTCACCATACATTTGGGCTCAG 3600  
3660 GTCCAGAAAGAGAGCTTACAGAGTCTAGGCTTTCACCATACATTTGGGCTCAG 3719  
3601 GGTTCAGAAAGAGGAGGAGGCAAGAAATTAAGAGTCTCAGAAAGAGAGTATCTAGAGAT 3660  
3720 GGTTCAGAAAGAGGAGGAGGCAAGAAATTAAGAGTCTCAGAAAGAGAGTATCTAGAGAT 3779  
3661 GAAGAGCTTCCCTGCTTCCACACCTGTTATTTGGTAAGTAAACATATACCTTCTCAG 3720  
3780 GAAGAGCTTCCCTGCTTCCACACCTGTTATTTGGTAAGTAAACATATACCTTCTCAG 3839  
3721 TCTACTAGGATAGCACCCGTTGCTACGAGTGTCTGTAAAGACAGAGAGAAATTA 3780  
3840 TCTACTAGGATAGCACCCGTTGCTACGAGTGTCTGTAAAGACAGAGAGAAATTA 3899  
3781 TTATCATTTGAAGATAGCTTAAATGACTGCAATACCGAGTAAATTTGGCAAGCATCT 3840  
3900 TTATCATTTGAAGATAGCTTAAATGACTGCAATACCGAGTAAATTTGGCAAGCATCT 3959  
3841 CAGGAATCATCTTATGATGAGGAAACAAATGTTGCTGATGCTGTTTCTTACAGTGC 3900  
3960 CAGGAATCATCTTATGATGAGGAAACAAATGTTGCTGATGCTGTTTCTTACAGTGC 4019  
3901 AGTGAATTTGAAGATAGCTTAAATGACTGCAATACCGAGTAAATTTGGTGGTCT 3960  
4020 AGTGAATTTGAAGATAGCTTAAATGACTGCAATACCGAGTAAATTTGGTGGTCT 4079  
3961 TCCAAACAAATGAGGATCATCTGTAAGGAGGAGTTGCTGATGAGTCAAGAAATG 4020  
4080 TCCAAACAAATGAGGATCATCTGTAAGGAGGAGTTGCTGATGAGTCAAGAAATG 4139  
4021 GTTTCAGATGATGAACAAAGAGAGAGGAGGCTTGAAGAAATTAATCAAGAGAGCAAGC 4080  
4140 GTTTCAGATGATGAACAAAGAGAGAGGAGGCTTGAAGAAATTAATCAAGAGAGCAAGC 4199  
4081 ATGGAATTTCAAACTTAGGTGAGAGGATCTGGGTGTGAGAGTGAACAAAGCTCTGAA 4140  
4200 ATGGAATTTCAAACTTAGGTGAGAGGATCTGGGTGTGAGAGTGAACAAAGCTCTGAA 4259  
4141 GACTGCTCAGGCTATCTCTCAGAGTGAATTTAACAATCAGCAGAGAGGATACCATG 4200  
4260 GACTGCTCAGGCTATCTCTCAGAGTGAATTTAACAATCAGCAGAGAGGATACCATG 4319  
4201 CAACATTAACCTGATTAAGGCTCCAGCAGAAATGGCTGAGACTACAACCTGTTAGAACG 4260  
4320 CAACATTAACCTGATTAAGGCTCCAGCAGAAATGGCTGAGACTACAACCTGTTAGAACG 4379  
4261 CATGGAGGAGGAGCTTCTAACAGCTACCTTCCATATAAGTACTCTTCTGCTTGAAG 4320  
4380 CATGGAGGAGGAGCTTCTAACAGCTACCTTCCATATAAGTACTCTTCTGCTTGAAG 4439  
4321 GACTGCGAAATTCAGAACAAAGCAGATCAGAAAAGAGCTATTAACCTCAGAAAAGT 4380  
4440 GACTGCGAAATTCAGAACAAAGCAGATCAGAAAAGAGCTATTAACCTCAGAAAAGT 4499  
4381 AGTGAATTAACCTTATAGCCAGAAATTCAGAAAGGAGGCTTCTGCTGAGAGTGGAGTCT 4440  
4500 AGTGAATTAACCTTATAGCCAGAAATTCAGAAAGGAGGCTTCTGCTGAGAGTGGAGTCT 4559  
4441 GCAGATAGTCTTACAGTAAATTAAGAAACAGAGAGTGAAGAGTCAATCCCTTTTAA 4500  
4560 GCAGATAGTCTTACAGTAAATTAAGAAACAGAGAGTGAAGAGTCAATCCCTTTTAA 4619  
4501 TGCCATCATTTAGATGATAGTGTGATACATGACAGTTGCTCTGAGAGTCTTCAAGATTA 4560

Db 4620 TGCCCATCATAGATAGATAGTGTACATGCACAGTTGCTCTGGAGTCTTCAGAAATAGA 4679  
 Oy 4561 AACTACCCATCTCAAGAGAGCTCATTAAGTGTGTGATGTGAGAGCAACAGCTGGAA 4620  
 Db 4680 AACTACCCATCTCAAGAGAGCTCATTAAGTGTGTGATGTGAGAGCAACAGCTGGAA 4739  
 Oy 4621 GAGTCTGGGCGACACAGATTTGACGAACATCTTACTTGCACAGGCAAGATCTAGAGGA 4680  
 Db 4740 GAGTCTGGGCGACACAGATTTGACGAACATCTTACTTGCACAGGCAAGATCTAGAGGA 4799  
 Oy 4681 ACCCTTACTTGAATCTGGAATCAGCTCTTCTGTGATGACCTGATATCTCTCT 4740  
 Db 4800 ACCCTTACTTGAATCTGGAATCAGCTCTTCTGTGATGACCTGATATCTCTCT 4859  
 Oy 4741 GAAGACAGAGCCCGACAGTCTGCTGTGGCAATACCATCTTCAACCTTGCAATG 4800  
 Db 4860 GAAGACAGAGCCCGACAGTCTGCTGTGGCAATACCATCTTCAACCTTGCAATG 4919  
 Oy 4801 AAAGTTCCCAATTTGAAGTTGCGAATCTGCCAGAGTCCAGCTCTGCTCATACT 4860  
 Db 4920 AAAGTTCCCAATTTGAAGTTGCGAATCTGCCAGAGTCCAGCTCTGCTCATACT 4979  
 Oy 4861 GATACTGCTGGGTATATGATCAATGGAAGAGTGTAGCAGGAGAGCAACGATTTGACA 4920  
 Db 4980 GATACTGCTGGGTATATGATCAATGGAAGAGTGTAGCAGGAGAGCAACGATTTGACA 5039  
 Oy 4921 GCTTCAACAGAAAGGCTCAACAAAAGATGCTCATGCTGTGCTGCTGACCCAGAA 4980  
 Db 5040 GCTTCAACAGAAAGGCTCAACAAAAGATGCTCATGCTGTGCTGCTGACCCAGAA 5099  
 Oy 4981 GAATTTATGCTGCTGATCAAGTTTGCAGAAAACACCATCATCTTAACTAATCTAAT 5040  
 Db 5100 GAATTTATGCTGCTGATCAAGTTTGCAGAAAACACCATCATCTTAACTAATCTAAT 5159  
 Oy 5041 ACTGAAGAGACATCATCTGTTTATGAAGAAAGATGCTGATGTTGTGTGACAGGACA 5100  
 Db 5160 ACTGAAGAGACATCATCTGTTTATGAAGAAAGATGCTGATGTTGTGTGACAGGACA 5219  
 Oy 5101 CTGAATATATTTCTAGGAATTCGGGAGGAAATGGTGGTTAGTATTTCTGGTGACC 5160  
 Db 5220 CTGAATATATTTCTAGGAATTCGGGAGGAAATGGTGGTTAGTATTTCTGGTGACC 5279  
 Oy 5161 CAGTCTATTTAAAGAAAAGAAATGCTGAATGAGCATGATTTTGAAGTCAAGAGAGATGTG 5220  
 Db 5280 CAGTCTATTTAAAGAAAAGAAATGCTGAATGAGCATGATTTTGAAGTCAAGAGAGATGTG 5339  
 Oy 5221 GTCATGGAAGAAACCAACAGTCCAAAGCGAGCAAGAAATCCAGAGACAGAAAGATC 5280  
 Db 5340 GTCATGGAAGAAACCAACAGTCCAAAGCGAGCAAGAAATCCAGAGACAGAAAGATC 5399  
 Oy 5281 TTCAGGGGGCTGAATCTGTTGCTATGAGGCGCTTCCACCAATGCGCCACAGATCAACTG 5340  
 Db 5400 TTCAGGGGGCTGAATCTGTTGCTATGAGGCGCTTCCACCAATGCGCCACAGATCAACTG 5459  
 Oy 5341 GAATGATGATGACAGCTGTGTGCTTCTGTGTGAGAGAGCTTTCATTCACCTT 5400  
 Db 5460 GAATGATGATGACAGCTGTGTGCTTCTGTGTGAGAGAGCTTTCATTCACCTT 5519  
 Oy 5401 GGCACAGGCTCCACCAATTTGTTGTTGCGAGCCAGATGCTTGGGCAAGAGACATTTG 5460  
 Db 5520 GGCACAGGCTCCACCAATTTGTTGTTGCGAGCCAGATGCTTGGGCAAGAGACATTTG 5579  
 Oy 5461 TTCATGCAATTTGGGCGAGATGTGTAGGACCTGTGTGACCCGAGAGTGGTGTGGAC 5520  
 Db 5580 TTCATGCAATTTGGGCGAGATGTGTAGGACCTGTGTGACCCGAGAGTGGTGTGGAC 5639  
 Oy 5521 AGTGTAGCACTTACCAAGTCCAGAGAGTGTGACACCTTACCTATATACCCAGATCCCCAC 5580  
 Db 5640 AGTGTAGCACTTACCAAGTCCAGAGAGTGTGACACCTTACCTATATACCCAGATCCCCAC 5699  
 Oy 5581 AGCCACTAC 5589  
 Db 5700 AGCCACTAC 5708

RESULT 13  
 AA287995  
 ID AA287995 standard; DNA: 5712 BP.  
 XX  
 AC AA287995:  
 XX  
 DT 06-JUN-2000 (first entry)  
 XX  
 DE BRCA1 gene sequence (GenBank Accn No: U14680).  
 XX  
 KW BRCA1; tumour; prostate cancer; cytostatic; antiproliferative;  
 KM gene therapy; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FH CDS 120..5711  
 FT /tag= a  
 FT /transl\_except- (pos: 1452..1454, aa:asp)  
 FT /transl\_except- (pos: 1482..1484, aa:asp)  
 FT /transl\_except- (pos: 1581..1583, aa:ser)  
 FT /transl\_except- (pos: 1899..1901, aa:glu)  
 FT /transl\_except- (pos: 1902..1904, aa:leu)  
 FT /transl\_except- (pos: 1905..1907, aa:glu)  
 FT /transl\_except- (pos: 1908..1910, aa:leu)  
 FT /transl\_except- (pos: 1911..1913, aa:asn)  
 FT /transl\_except- (pos: 1914..1916, aa:ile)  
 FT /transl\_except- (pos: 1917..1919, aa:met)  
 FT /transl\_except- (pos: 4476..4478, aa:val)  
 FT /transl\_except- (pos: 4479..4481, aa:leu)  
 FT /transl\_except- (pos: 4482..4484, aa:glu)  
 FT /transl\_except- (pos: 4533..4535, aa:tyr)  
 FT /transl\_except- (pos: 4689..4691, aa:pro)  
 FT /note= "Xaa- unknown"  
 XX  
 PN CA2218197-A1.  
 XX  
 PD 05-JUN-1999.  
 XX  
 PE 12-DEC-1997; 97CA-2218197.  
 XX  
 PR 05-DEC-1997; 97US-0986106.  
 XX  
 PA (UNIV ) UNIV WASHINGTON.  
 PA (UYVA-) UNIV VANDERBILT.  
 PI Robinson-Benion CL, Thompson ME, Holt JT, Jensen RA, Steiner MS;  
 PI King M;  
 DR WPI: 2000-238071/21.  
 DR P-PSDB; AAT77818.  
 XX  
 PT New method of treatment and suppression of prostate cancer comprises  
 PT using the BRCA family of genes to decrease the growth rate of the tumor  
 PT  
 PS Claim 11: Page 98-104; 16pp; English.  
 XX  
 CC The invention relates to a method for suppressing the growth of a  
 CC prostate tumour in a mammal that comprises introducing to the tumour a  
 CC vector comprising a nucleic acid sequence encoding a BRCA family gene  
 CC product operatively linked to a promoter, where production of the BRCA  
 CC family gene product results in a decrease in the growth rate of the  
 CC tumour. The methods are used to suppress the growth of and also to treat  
 CC prostate tumour in a mammal where the tumour is gene-linked hereditary  
 CC prostate cancer or sporadic prostate cancer. The present sequence  
 CC represents a BRCA1 gene sequence (GenBank Accn No: U14680) that can be  
 CC administered for suppressing tumour.  
 XX  
 SQ Sequence 5712 BP; 1956 A; 1099 C; 1274 G; 1383 T; 0 other;

Query Match 99.9%; Score 5585.8; DB 21; Length 5712;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 5587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ARGAGATTATCTGCTCTTGGCGTTGAGAGATCAAAATGTCTATATGCTATGCGAAA 60  
DB 120 ARGAGATTATCTGCTCTTGGCGTTGAGAGATCAAAATGTCTATATGCTATGCGAAA 179  
QY 61 AACTGAGATGCCCTGCTGCTGAGTGTATCAAGAACCTGCTCCCAAAAGTGTGAC 120  
DB 180 AACTGAGATGCCCTGCTGCTGAGTGTATCAAGAACCTGCTCCCAAAAGTGTGAC 239  
QY 121 CACATATTTGCAAAATTTGCAAGTGTGAACTTCTCAACGAGAAAGAGGCTTCCACAG 180  
DB 240 CACATATTTGCAAAATTTGCAAGTGTGAACTTCTCAACGAGAAAGAGGCTTCCACAG 239  
QY 181 TGTCTTTTATGTAAGATGATATACCAAAAGAGGCTTCAAGAAAGTACGAGATTAGT 240  
DB 300 TGTCTTTTATGTAAGATGATATACCAAAAGAGGCTTCAAGAAAGTACGAGATTAGT 359  
QY 241 CAACCTGTGAAGAGCTATTTGAAATCATTTGTCTTCACTTGAACAGGTTTGGAG 300  
DB 360 CAACCTGTGAAGAGCTATTTGAAATCATTTGTCTTCACTTGAACAGGTTTGGAG 419  
QY 301 TATGCAAAAGAGCTATATTTTGCAGAAAGGAAATTAATCTCTGACATCTTAAAGAT 360  
DB 420 TATGCAAAAGAGCTATATTTTGCAGAAAGGAAATTAATCTCTGACATCTTAAAGAT 479  
QY 361 GAAGTTTCTATCTCAAAAGTATGGCTACAGAAACCGTCCAAAGACTTCTACAGAGT 420  
DB 480 GAAGTTTCTATCTCAAAAGTATGGCTACAGAAACCGTCCAAAGACTTCTACAGAGT 539  
QY 421 GAACCGAAATCCCTCTGAGAAACAGCTCAGTGTCCAGCTCTAATCTTGGGA 480  
DB 540 GAACCGAAATCCCTCTGAGAAACAGCTCAGTGTCCAGCTCTAATCTTGGGA 539  
QY 481 ACTGTGAGAACTCTGAGAAACAGCGGATCAACCTCAAAAGAGCTGTCTACAT 540  
DB 600 ACTGTGAGAACTCTGAGAAACAGCGGATCAACCTCAAAAGAGCTGTCTACAT 659  
QY 541 GAATGGGATCTGATCTTCTGAAAGTACCCTTAATTAAGCAATTAATGCAATGCGGA 600  
DB 660 GAATGGGATCTGATCTTCTGAAAGTACCCTTAATTAAGCAATTAATGCAATGCGGA 719  
QY 601 GATCAAGATTTGTAACAATCACCCCTCAAGGAACAGGATGAATTCAGTTGGATTC 660  
DB 720 GATCAAGATTTGTAACAATCACCCCTCAAGGAACAGGATGAATTCAGTTGGATTC 779  
QY 661 GCAAAAAAGGCTGCTGTGAATTTTCTGAGAGGATGTAACAAACTGCAATCATCA 720  
DB 780 GCAAAAAAGGCTGCTGTGAATTTTCTGAGAGGATGTAACAAACTGCAATCATCA 839  
QY 721 CCCAGTAATATGATTTGAAACACACTGAGAAAGCTGACGTGAGAGGATCCAGAAA 780  
DB 840 CCCAGTAATATGATTTGAAACACACTGAGAAAGCTGACGTGAGAGGATCCAGAAA 839  
QY 781 TATCAGGATGTTCTGTTCAAACTGCATGTGAGGACCATGTGGGCAAAATACGTG 840  
DB 900 TATCAGGATGTTCTGTTCAAACTGCATGTGAGGACCATGTGGGCAAAATACGTG 959  
QY 841 AGCTCATTAACAGATGAGAACAGCACTTATTAATCTCAATAAGACAGAAATGTAGAA 900  
DB 960 AGCTCATTAACAGATGAGAACAGCACTTATTAATCTCAATAAGACAGAAATGTAGAA 1019  
QY 901 AAGGCTGAATTTCTGTATATTAAGCAACAGCTGCTTACGAAGAGCAATTAAGAA 960  
DB 1020 AAGGCTGAATTTCTGTATATTAAGCAACAGCTGCTTACGAAGAGCAATTAAGAA 1079  
QY 961 TGGGCTGGAAGTAAAGCAATGTATGATGAGGAGCTCCAGACAGAGAAAAAGGTA 1020  
DB 1080 TGGGCTGGAAGTAAAGCAATGTATGATGAGGAGCTCCAGACAGAGAAAAAGGTA 1139  
QY 1021 GATCTGAATGTGATCCCTGTGTGAGAGAAAAAGATGAATTAAGCAAACTGCCATGC 1080

DB 1140 GATCTGAATGTGATCCCTGTGTGAGAGAAAAAGATGAATTAAGCAAACTGCCATGC 1199  
QY 1081 TCAGAGAAATCTGAGATATCTGAAGATGTTCTTGGATTAACACTTAATAGACATTCAG 1140  
DB 1200 TCAGAGAAATCTGAGATATCTGAAGATGTTCTTGGATTAACACTTAATAGACATTCAG 1259  
QY 1141 AAGTTAATGATGTTTCTCAGAAAGTATGATGTTAGTCTGTATGCTACATGAT 1200  
DB 1260 AAGTTAATGATGTTTCTCAGAAAGTATGATGTTAGTCTGTATGCTACATGAT 1319  
QY 1201 GGGAGCTGTAATCAAAATGCAAAAGTACGTATGTATGAGACCTTCTAAATGAGTAGAT 1260  
DB 1320 GGGAGCTGTAATCAAAATGCAAAAGTACGTATGTATGAGACCTTCTAAATGAGTAGAT 1379  
QY 1261 GAATATTTCTGTTCTTCAAGAAATTAAGATTAATGCTGAGGAGTCTCTAGAGCTTTA 1320  
DB 1380 GAATATTTCTGTTCTTCAAGAAATTAAGATTAATGCTGAGGAGTCTCTAGAGCTTTA 1439  
QY 1321 ATATGTAAAGTGAAGAGTTCACCTCAAAATCAGTAGAGATTAATTAAGCAAAATA 1380  
DB 1440 ATATGTAAAGTGAAGAGTTCACCTCAAAATCAGTAGAGATTAATTAAGCAAAATA 1499  
QY 1381 TTTGGAAAACTATCGAAGAGGCAAGGCTCCCAACTTAAGCCATGTACGAAAT 1440  
DB 1500 TTTGGAAAACTATCGAAGAGGCAAGGCTCCCAACTTAAGCCATGTACGAAAT 1559  
QY 1441 CTATATTAAGGACATTTTGTACTGAGCCCAATATTAAGAGGTCCTCCACAAAT 1500  
DB 1560 CTATATTAAGGACATTTTGTACTGAGCCCAATATTAAGAGGTCCTCCACAAAT 1619  
QY 1501 AAATTAAGGCTTAAGAGAGACCTACATCAGGCTTCATCTGAGATTTTATCAAGAAA 1560  
DB 1620 AAATTAAGGCTTAAGAGAGACCTACATCAGGCTTCATCTGAGATTTTATCAAGAAA 1679  
QY 1561 GCAGATTTGGCACTTCAAAAGAGCTCTGAATGATTAATCAGGAACTAACCAACGAG 1620  
DB 1680 GCAGATTTGGCACTTCAAAAGAGCTCTGAATGATTAATCAGGAACTAACCAACGAG 1739  
QY 1621 CAGATGCTCAATGATGATTAATTAATGATGCTATGATTAATTAAGAAAGGAT 1680  
DB 1740 CAGATGCTCAATGATGATTAATTAATGATGCTATGATTAATTAAGAAAGGAT 1799  
QY 1681 TCTATTCAGATGAGAAAAATCTTAACCCATTAATGATCACTGCAAAAGATTCGCTTC 1740  
DB 1800 TCTATTCAGATGAGAAAAATCTTAACCCATTAATGATCACTGCAAAAGATTCGCTTC 1859  
QY 1741 AAAACGAAAGCTGAACTATTAAGCAGCAGTATTAAGCAATTAAGCAATTAATATC 1800  
DB 1860 AAAACGAAAGCTGAACTATTAAGCAGCAGTATTAAGCAATTAAGCAATTAATATC 1919  
QY 1801 CACATTTCAAAAGCACCCTTAAGAAATAGGCTGAGAGAGAGCTTCTACAGGCAATTT 1860  
DB 1920 CACATTTCAAAAGCACCCTTAAGAAATAGGCTGAGAGAGAGCTTCTACAGGCAATTT 1979  
QY 1861 CATGCGCTGAATCTAGTACGATTAAGCAATTAAGCAATTAAGCAATTAAGCAAT 1920  
DB 1980 CATGCGCTGAATCTAGTACGATTAAGCAATTAAGCAATTAAGCAATTAAGCAAT 2039  
QY 1921 ATTGATAGTTGTTCTAGCAGTGAAGATTAAGAAAAAAAGTACAAACCAATGCCAGTC 1980  
DB 2040 ATTGATAGTTGTTCTAGCAGTGAAGATTAAGAAAAAAAGTACAAACCAATGCCAGTC 2099  
QY 1981 AGGACAGCAGAAACCTACACATCATGAGAGGTAAGAAACCTGCAACTGAGGCAAGAG 2040  
DB 2100 AGGACAGCAGAAACCTACACATCATGAGAGGTAAGAAACCTGCAACTGAGGCAAGAG 2159  
QY 2041 AGTAACAGCAAAATGAACAGACAACTTAAGAAAGCAATGAGATCTTTCCAGAGCTG 2100  
DB 2160 AGTAACAGCAAAATGAACAGACAACTTAAGAAAGCAATGAGATCTTTCCAGAGCTG 2219  
QY 2101 AAGTTAACAATGACCTGTTCTTTTACTAAGTGTCAATTAACAGTGAATTAAGAA 2160



```
Db 2220 AAGTTAACAAATGCACCTGGTTCTTTTACTAAGTGTCAAAATACGAGTGAACCTTAAAGAA 2279
QY 2161 TTGTCAATCCAGCTTCACAAGAGAGAAAAAAGAGAACTGAAACAGTTAAAGTG 2220
Db 2280 TTTGTCAATCCTAGCTTCACAAGAGAGAAAAAAGAGAACTGAAACAGTTAAAGTG 2339
QY 2221 TCTAATAATGCTGAAGAGACCCCAAGATCTCATGTTAAGTGAGAAAGGGTTTGGCAACT 2280
Db 2340 TCTAATAATGCTGAAGAGACCCCAAGATCTCATGTTAAGTGAGAAAGGGTTTGGCAACT 2399
QY 2281 GAAAGATCTGTAGAGAGAGAGATTTTCATTGGTACCTGGTACTGATTATGGCACTCG 2340
Db 2400 GAAAGATCTGTAGAGAGAGAGATTTTCATTGGTACCTGGTACTGATTATGGCACTCG 2459
QY 2341 GAAAGTATCTGCTTCTGAAAGTATGACACTTAGGAGAAAGGCAAAAAACAGAACCAATAA 2400
Db 2460 GAAAGTATCTGCTTCTGAAAGTATGACACTTAGGAGAAAGGCAAAAAACAGAACCAATAA 2519
QY 2401 TGTGTAGTCAAGTGTGACAGCATTTTGAACCCCAAGGACATAATTCTATGGTTGTCCAAA 2460
Db 2520 TGTGTAGTCAAGTGTGACAGCATTTTGAACCCCAAGGACATAATTCTATGGTTGTCCAAA 2579
QY 2461 GATATATATAAATGACACAGAAAGGCTTTAAGTATCCATTGGGACATGAACTTAACCACT 2520
Db 2580 GATATATATAAATGACACAGAAAGGCTTTAAGTATCCATTGGGACATGAACTTAACCACT 2639
QY 2521 CGGGAACAACAGCATGGAATGGAAGAAAGTGAATGATGCTCAGTATTGGAGAAATACA 2580
Db 2640 CGGGAACAACAGCATGGAATGGAAGAAAGTGAATGATGCTCAGTATTGGAGAAATACA 2699
QY 2581 TTCAAGGTTTCAAGGCGCAGTCATTTTGCTGTCTGTTTCAATCCAGAAATCAGAAAG 2640
Db 2700 TTCAAGGTTTCAAGGCGCAGTCATTTTGCTGTCTGTTTCAATCCAGAAATCAGAAAG 2759
QY 2641 GAATGTGCAATCTCTGCCCACCTGGGTCCTTAAAGAAACAAAGTCCAAAGCACT 2700
Db 2760 GAATGTGCAATCTCTGCCCACCTGGGTCCTTAAAGAAACAAAGTCCAAAGCACT 2819
QY 2701 TTTGATGTGAACAAGAAAGAAATCAAGGAAAGAAATGAGTCTAATATCAAGCCTGTA 2760
Db 2820 TTTGATGTGAACAAGAAAGAAATCAAGGAAAGAAATGAGTCTAATATCAAGCCTGTA 2879
QY 2761 CAGACAGTTAATATCAGTCGAGGCTTCTGTGTGGTGGTCAGAAAGATTAAGCAGTTGAT 2820
Db 2880 CAGACAGTTAATATCAGTCGAGGCTTCTGTGTGGTGGTCAGAAAGATTAAGCAGTTGAT 2939
QY 2821 AATGCAATGTAGTATCAAGAGGAGCTAGTTTGTCTATCACTCTCAGTTCAGAGGC 2880
Db 2940 AATGCAATGTAGTATCAAGAGGAGCTAGTTTGTCTATCACTCTCAGTTCAGAGGC 2999
QY 2881 AACGAACTGAGACTTACTCCAAATAAACATGAGACTTTTACAAACCCCATATGCTATA 2940
Db 3000 AACGAACTGAGACTTACTCCAAATAAACATGAGACTTTTACAAACCCCATATGCTATA 3059
QY 2941 CCACCACTTTTCCCATCAAGTCATTTGTTAAACTAAATGTAGAAAAATCTGCTAGAG 3000
Db 3060 CCACCACTTTTCCCATCAAGTCATTTGTTAAACTAAATGTAGAAAAATCTGCTAGAG 3119
QY 3001 GAAAACTTGAAGAACATCAATGTCACCTGAAAGAGAAATGGGAAATGAGAACATTCOA 3060
Db 3120 GAAAACTTGAAGAACATCAATGTCACCTGAAAGAGAAATGGGAAATGAGAACATTCOA 3179
QY 3061 AGTACAGTGAAGACATTAAGCGGTATTAACATTAAGAGAAATGTTTAAAGAGCAGAC 3120
Db 3180 AGTACAGTGAAGACATTAAGCGGTATTAACATTAAGAGAAATGTTTAAAGAGCAGAC 3239
QY 3121 TCAAGCAATATTAATGAAGTAGTTCAGTACTAATGAAGTGGGCTCCAGTATTATGAA 3180
Db 3240 TCAAGCAATATTAATGAAGTAGTTCAGTACTAATGAAGTGGGCTCCAGTATTATGAA 3299
QY 3181 ATAGGTTCCAGTGAAGAAACATTCAGCAGAACTAGTGAAGAAACAGAGGCGCAAAATTG 3240
Db 3300 ATAGGTTCCAGTGAAGAAACATTCAGCAGAACTAGTGAAGAAACAGAGGCGCAAAATTG 3359
QY 3241 AATGCTATGCTAGATTAAGGGGTTTGGCAACCTGAGGCTATTAACAAAGCTTCTCTGA 3300
Db 3360 AATGCTATGCTAGATTAAGGGGTTTGGCAACCTGAGGCTATTAACAAAGCTTCTCTGA 3419
QY 3301 AGTAAATTTGAACATCTGAAATTAAGAAACAGAAATATGAAGAGTGTTCAGACTGT 3360
Db 3420 AGTAAATTTGAACATCTGAAATTAAGAAACAGAAATATGAAGAGTGTTCAGACTGT 3479
QY 3361 AATACAGATTTCTCCAAATGATTTTCAGTAACTAGAACAGCCATAGGAGTAGT 3420
Db 3480 AATACAGATTTCTCCAAATGATTTTCAGTAACTAGAACAGCCATAGGAGTAGT 3539
QY 3421 CATGCACTCAGGTTTGTCTGAGACACCTGATGACCTGTTAGATGATGAAATTAAG 3480
Db 3540 CATGCACTCAGGTTTGTCTGAGACACCTGATGACCTGTTAGATGATGAAATTAAG 3599
QY 3481 GAAGATCTAGTTTGTCTGAAATATGACATTAAGAAAGTGTGCTGTTTAAAGAAAGC 3540
Db 3600 GAAGATCTAGTTTGTCTGAAATATGACATTAAGAAAGTGTGCTGTTTAAAGAAAGC 3659
QY 3541 GTCCAGAAAGAGAGCTTAGCAGAGTCTAGCCCTTACCCATACACATTTGGCTCAG 3600
Db 3660 GTCCAGAAAGAGAGCTTAGCAGAGTCTAGCCCTTACCCATACACATTTGGCTCAG 3719
QY 3601 GGTACCCGAAAGAGGCGCCAGAAATTAAGAGTCTCAGAAAGAACTTATCTAGAGAT 3660
Db 3720 GGTACCCGAAAGAGGCGCCAGAAATTAAGAGTCTCAGAAAGAACTTATCTAGAGAT 3779
QY 3661 GAAGAGCTTCCCTGCTCCAAACACTGTTTATTTGGTAAAGTAAACATATACCTTCAG 3720
Db 3780 GAAGAGCTTCCCTGCTCCAAACACTGTTTATTTGGTAAAGTAAACATATACCTTCAG 3839
QY 3721 TCTACTAGGCAATAGACCGTGTCTACGAGTGTCTGTAAAGAACACAGAGAGAAATTA 3780
Db 3840 TCTACTAGGCAATAGACCGTGTCTACGAGTGTCTGTAAAGAACACAGAGAGAAATTA 3899
QY 3781 TTATCATTTGAAGAAATAGCTTAAATGACTGAGTAACCAAGTAAATTTGGCAAGGACTCT 3840
Db 3900 TTATCATTTGAAGAAATAGCTTAAATGACTGAGTAACCAAGTAAATTTGGCAAGGACTCT 3959
QY 3841 CAGGAACATCACTTATGAGAGAGAAACAAATGTTCTGTACTGTGTTTCTTCACAGTGC 3900
Db 3960 CAGGAACATCACTTATGAGAGAGAAACAAATGTTCTGTACTGTGTTTCTTCACAGTGC 4019
QY 3901 AGTGAATTTGAAGACTTACTGCAATATCAAAACACCAGATCTTCTTCTGATGATGCTCT 3960
Db 4020 AGTGAATTTGAAGACTTACTGCAATATCAAAACACCAGATCTTCTTCTGATGATGCTCT 4079
QY 3961 TCCAAAACAAATGAGGATCAGTCTGAAAGCCAGGAGTTGGTCTGAGTACAGAAATTG 4020
Db 4080 TCCAAAACAAATGAGGATCAGTCTGAAAGCCAGGAGTTGGTCTGAGTACAGAAATTG 4139
QY 4021 GTTTCAGATGATGAAGAAAGAGAGAAACGGGCTTGGAGAAATAATATCAAGAGAGCAAGC 4080
Db 4140 GTTTCAGATGATGAAGAAAGAGAGAAACGGGCTTGGAGAAATAATATCAAGAGAGCAAGC 4199
QY 4081 AATGATTTCAAACTTGTGTGAAGCAGCATCTGGGTGTGAGATGAAACAGGCTCTCGAA 4140
Db 4200 AATGATTTCAAACTTGTGTGAAGCAGCATCTGGGTGTGAGATGAAACAGGCTCTCGAA 4259
QY 4141 GACTGCTGAGGCTATCTCTCAGAGTACATTTTAAACACTCAGCAGAGAGATACCAT 4200
Db 4260 GACTGCTGAGGCTATCTCTCAGAGTACATTTTAAACACTCAGCAGAGAGATACCAT 4319
QY 4201 CAACATTAACCTGATTAAGCTCCAGCAGGAAATGCTGTAAGTGAAGCTGTGTTAAGACAG 4260
Db 4320 CAACATTAACCTGATTAAGCTCCAGCAGGAAATGCTGTAAGTGAAGCTGTGTTAAGACAG 4379
QY 4261 CATGGAGGCGAGGCTTCTTAACAGTACCTCTCCATTAATGAGTCTGTGGCCCTGAG 4320
Db 4380 CATGGAGGCGAGGCTTCTTAACAGTACCTCTCTCCATTAATGAGTCTGTGGCCCTGAG 4439
```



```

QY 4321 GACCTGGGAATTCAGAACAAAGCAGATCAGAAAAGCACTATTAACTTCAGAGAAAGT 4380
    |||
Db 4440 GACCTGGGAATTCAGAACAAAGCAGATCAGAAAAGCACTATTAACTTCAGAGAAAGT 4499
QY 4381 AGTGAATACCTTTAAGCCAGAAATCCAGAAAGGCTTTCTCTGACAAAGTTTGAGGTCT 4440
    |||
Db 4500 AGTGAATACCTTTAAGCCAGAAATCCAGAAAGGCTTTCTCTGACAAAGTTTGAGGTCT 4559
QY 4441 GCAGATAGTTCTACAGATAAATAAAGAACAGAGAGATGGAAGGTATCCCTTTAAA 4500
    |||
Db 4560 GCAGATAGTTCTACAGATAAATAAAGAACAGAGAGATGGAAGGTATCCCTTTAAA 4619
QY 4501 TGGCCATCATTTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4560
    |||
Db 4620 TGGCCATCATTTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4679
QY 4561 AACTACCATCTCAAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 4620
    |||
Db 4680 AACTACCATCTCAAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 4739
QY 4621 GAGTCTGGGCGACACGATTTGAGGAAACATCTTACTTGGCAAGGCAAGATCTAGAGGA 4680
    |||
Db 4740 GAGTCTGGGCGACACGATTTGAGGAAACATCTTACTTGGCAAGGCAAGATCTAGAGGA 4799
QY 4681 ACCCTTACCTGGAATCTGGAATCAGCCTCTCTGATGATGATGATGATGATGATGATGAT 4740
    |||
Db 4800 ACCCTTACCTGGAATCTGGAATCAGCCTCTCTGATGATGATGATGATGATGATGATGAT 4859
QY 4741 GAAGACAGAGCCCGACAGTACGCTCGTGTGGCAACATACCATCTTCACCTTCGATTG 4800
    |||
Db 4860 GAAGACAGAGCCCGACAGTACGCTCGTGTGGCAACATACCATCTTCACCTTCGATTG 4919
QY 4801 AAGGTTCCCAATTTGAAAGTTGCAAGATGCGCCAGAGTCCAGTCTGCTCATCTACT 4860
    |||
Db 4920 AAGGTTCCCAATTTGAAAGTTGCAAGATGCGCCAGAGTCCAGTCTGCTCATCTACT 4979
QY 4861 GATACGCTGGTATATGCAATGCAAGAAAGTGTGAGCAGGAGCAAGCCAGATTGACA 4920
    |||
Db 4980 GATACGCTGGTATATGCAATGCAAGAAAGTGTGAGCAGGAGCAAGCCAGATTGACA 5039
QY 4921 GCTTCAACAGAAAGGCTCAACAAAGATGCTCATGCTGCTGCTGACCCAGAA 4980
    |||
Db 5040 GCTTCAACAGAAAGGCTCAACAAAGATGCTCATGCTGCTGCTGACCCAGAA 5099
QY 4981 GAATTTATGCTGCTGATGCAAGTTTGCAGAAAACACCATCTTAACTAATCTAAT 5040
    |||
Db 5100 GAATTTATGCTGCTGATGCAAGTTTGCAGAAAACACCATCTTAACTAATCTAAT 5159
QY 5041 ACTGAAGAGACTACTCATGTTGTTATGAAAACAGATGCTGAGTTTGTGTGAACGACA 5100
    |||
Db 5160 ACTGAAGAGACTACTCATGTTGTTATGAAAACAGATGCTGAGTTTGTGTGAACGACA 5219
QY 5101 CTGAATATTTTATGGAATTTGGGAGGAAATGGGTAGTTAGCTATTTCTGGGTGACC 5160
    |||
Db 5220 CTGAATATTTTATGGAATTTGGGAGGAAATGGGTAGTTAGCTATTTCTGGGTGACC 5279
QY 5161 CACTCATTTAAAGAAAGAAAATGCTGAATGAGCATGATTTTGAAGTCAGAGAGATG 5220
    |||
Db 5280 CACTCATTTAAAGAAAGAAAATGCTGAATGAGCATGATTTTGAAGTCAGAGAGATG 5339
QY 5221 GTCGAATGGAAGAACCAAGGTTCCAAAGCGAGAGAAATCCAGACAGAAAGATC 5280
    |||
Db 5340 GTCGAATGGAAGAACCAAGGTTCCAAAGCGAGAGAAATCCAGACAGAAAGATC 5399
QY 5281 TTGAGGGGGCTAGAAATCTGTTCTATGGCCCTTACCAACATGCCCACAGATCAACTG 5340
    |||
Db 5400 TTGAGGGGGCTAGAAATCTGTTCTATGGCCCTTACCAACATGCCCACAGATCAACTG 5459
QY 5341 GAATGATGGTACAGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5400
    |||
Db 5460 GAATGATGGTACAGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5519
QY 5401 GGCACAGGTGTCCACCAATTGTGTGTGCAAGCCAGATGCTTGAGCAGAGCAATGCG 5460

```

```

Db 5520 GGCACAGGTGTCCACCAATTGTGTGTGCGAGCAGATGCTGAGACAGAGACAAATGCG 5579
QY 5461 TTCCATGCATTTGGGCGAGATGTGTGAGGACACCTGTGTGACCCGAGAGTGGTGTGAC 5520
    |||
Db 5580 TTCCATGCATTTGGGCGAGATGTGTGAGGACACCTGTGTGACCCGAGAGTGGTGTGAC 5639
QY 5521 AGTGTACACTCTACAGTCCAGAGAGTGTGAGACACCTACCTGATATACCCAGATCCCCAC 5580
    |||
Db 5640 AGTGTACACTCTACAGTCCAGAGAGTGTGAGACACCTACCTGATATACCCAGATCCCCAC 5699
QY 5581 AGCCACTAC 5589
    |||
Db 5700 AGCCACTAC 5708

RESULT 14
AAT32601
ID AAT32601 standard; cDNA; 5914 BP.
XX
AC AAT32601;
XX
DT 19-NOV-1996 (first entry)
XX
DE BRCA1, breast and ovarian cancer susceptibility gene.
XX
KW BRCA1: breast cancer; ovary cancer; predisposing gene; diagnosis;
KW susceptibility gene; prognosis; gene therapy; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 120..5711
FT /*tag= a
FT /product= BRCA1 protein
XX
PN MO9605308-A1.
XX
PD 22-FEB-1996.
XX
PF 11-AUG-1995; 95WO-US10220.
XX
PR 07-JUN-1995; 95US-0488011.
PR 12-AUG-1994; 94US-0289221.
PR 02-SEP-1994; 94US-0300266.
PR 16-SEP-1994; 94US-0308104.
PR 29-NOV-1994; 94US-0348824.
PR 24-MAR-1995; 95US-0409305.
PR 07-JUN-1995; 95US-0483554.
PR 07-JUN-1995; 95US-0487002.
XX
PA (MYRI-) MYRIAD GENETICS INC.
PA (UTAH ) UNIV UTAH RES FOUND.
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Futreal PA, Goldgar DE, Harshtman KD, Kamb A, Miki Y,
PI Shattuck-Eidens DM, Skolnick MH, Swenson J, Tavtigian SV,
PI Wiseman RW;
XX
DR MPI; 1996-139704/14.
DR P-PDB; AAR97128.
XX
PT New method for diagnosing a predisposition to breast and ovarian
PT cancer - by detecting a germline alteration in the BRCA1 gene or
PT gene regulatory sequence; for gene therapy and to screen for drugs
XX
PS Claim 4; Page 108-117; 200pp; English.
XX
CC This is the nucleotide sequence of the breast and ovarian cancer
CC susceptibility gene, BRCA1. Four kindred families provided genetic
CC evidence for localisation of BRCA1 to a sufficiently small region for
CC the application of positional cloning strategies. A detailed map of
CC transcripts was developed for the region of 17q21 between D17S1321 and

```

CC D1/S1324. A combination of sequences obtd. from cDNA clones,  
CC hybrid-selected sequences and PCR prods. allowed construction of a  
CC composite full-length BRCA1 cDNA (see AAT32612 for genomic sequence).  
CC The isolated cDNA is used in methods for either diagnosis of the  
CC predisposition to cancer (famil. breast and ovarian cancer), or for the  
CC diagnosis or prognosis of cancer, and also in gene-based therapies  
CC directed at cancer cells.  
XX

SQ Sequence 5914 BP; 2006 A; 1156 C; 1316 G; 1436 T; 0 other;

Query Match 99.9%; Score 5585.8; DB 17; Length 5914;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 5587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 ATGATTTATCTGCTCTGCGCTTGAAGAATGCAAAATGCTATATGCTATGACAGAA 60  
120 ATGATTTATCTGCTCTGCGCTTGAAGAATGCAAAATGCTATATGCTATGACAGAA 179  
61 ATCTTGAAGTGTCCATCTGCTGTGAGTATCAAGAACCTGCTCCACAAAGTGTAC 120  
180 ATCTTGAAGTGTCCATCTGCTGTGAGTATCAAGAACCTGCTCCACAAAGTGTAC 239  
121 CACATATTTGCAAAATTTGCAATGCTGCAAACTCTGACCCAGAAAGAGGCTTCACAG 180  
240 CACATATTTGCAAAATTTGCAATGCTGCAAACTCTGACCCAGAAAGAGGCTTCACAG 299  
181 TGTCTTTATGTAGATGATATTAACCAAGAGGCTTACAAAGAAAGTACGAGATTAGT 240  
300 TGTCTTTATGTAGATGATATTAACCAAGAGGCTTACAAAGAAAGTACGAGATTAGT 359  
241 CACTTTGTTGAAGAGCTATTGAAATCATTTGTCTTTTCAGCTTGACACAGGTTTGGAG 300  
360 CACTTTGTTGAAGAGCTATTGAAATCATTTGTCTTTTCAGCTTGACACAGGTTTGGAG 419  
301 TATGCAACAGCTATTAATTTTGGCAAAAGAAATTAACCTCTGACCACTTAAAGAT 360  
420 TATGCAACAGCTATTAATTTTGGCAAAAGAAATTAACCTCTGACCACTTAAAGAT 479  
361 GAAGTTTCTATCATCCAAAGTATGGGCTACAGAAACCGTGCCAAAGAGCTTCTACAGAGT 420  
480 GAAGTTTCTATCATCCAAAGTATGGGCTACAGAAACCGTGCCAAAGAGCTTCTACAGAGT 539  
421 GAACCCGAAATTCCTTCTTGACAGAAACCGAGTCTAGTGTCCAACTCTTAACTTTGA 480  
540 GAACCCGAAATTCCTTCTTGACAGAAACCGAGTCTAGTGTCCAACTCTTAACTTTGA 539  
481 ACTGTGAGAACTCTGAGCAAAAGAGGCGGATACAACTCAAAAGAGCTGTCTACAT 540  
600 ACTGTGAGAACTCTGAGCAAAAGAGGCGGATACAACTCAAAAGAGCTGTCTACAT 659  
541 GAATTTGGGATCTGATTCCTGTAAGATACCGTTAATAGGCAACTTATGCACTGTGGGA 600  
660 GAATTTGGGATCTGATTCCTGTAAGATACCGTTAATAGGCAACTTATGCACTGTGGGA 719  
601 GATCAAGAATTTGTACAAATTCACCCCTCAAGGAGGAGGATGAATTCAGTTTGAATTC 660  
720 GATCAAGAATTTGTACAAATTCACCCCTCAAGGAGGAGGATGAATTCAGTTTGAATTC 779  
661 GCAAAAAGAGGCTGTGTAATTTTCTGAGAGGATGTAACAATCTGCAACTCATCA 720  
780 GCAAAAAGAGGCTGTGTAATTTTCTGAGAGGATGTAACAATCTGCAACTCATCA 839  
721 CCCAGTAATATGATTTGAACACCACTGAGAGCGTGCAGCTGAGAGGATCCAGAAAG 780  
840 CCCAGTAATATGATTTGAACACCACTGAGAGCGTGCAGCTGAGAGGATCCAGAAAG 839  
781 TATCAGGGTAGTCTGTTCAAACTTGATGTGAGCCATGTGGCAACAATACTCATGCC 840  
900 TATCAGGGTAGTCTGTTCAAACTTGATGTGAGCCATGTGGCAACAATACTCATGCC 959  
841 AGCTATTACAGCATGAGACAGACAGCTTTATCTACTAAGACAGATGAATGTAGAA 900  
960 AGCTATTACAGCATGAGACAGACAGCTTTATCTACTAAGACAGATGAATGTAGAA 1019

901 AAGCTGAATTTCTGTATATAAAGCAACAGCCGTGCTTAGCAAGAGGCCATTAACAGA 960  
1020 AAGCTGAATTTCTGTATATAAAGCAACAGCCGTGCTTAGCAAGAGGCCATTAACAGA 1079  
961 TGGGCTGGAAGTAGAAGCAATGATATGATAGCCGACCTCCACAGACAGAAAAAGTA 1020  
1080 TGGGCTGGAAGTAGAAGCAATGATATGATAGCCGACCTCCACAGACAGAAAAAGTA 1139  
1021 GATCTGAATGCTGATCCCTGTGTGAGAGAAAGATGGAATTAACACAACTGCCATGC 1080  
1140 GATCTGAATGCTGATCCCTGTGTGAGAGAAAGATGGAATTAACACAACTGCCATGC 1199  
1081 TCAGAAATCTTGAAGATCTGAGATGCTTCTGATTAACACTAAATAGCAGATTCAG 1140  
1200 TCAGAAATCTTGAAGATCTGAGATGCTTCTGATTAACACTAAATAGCAGATTCAG 1259  
1141 AAGTTAATGATGCTTTTCCAGAGTATGATGCTTATGCTGATGACTCAGATGAT 1200  
1260 AAGTTAATGATGCTTTTCCAGAGTATGATGCTTATGCTGATGACTCAGATGAT 1319  
1201 GGGAGTCTGAATCAATGCAAGTATGCTGATGCTTATGCTGATGACTCAGATGAT 1260  
1320 GGGAGTCTGAATCAATGCAAGTATGCTGATGCTTATGCTGATGACTCAGATGAT 1379  
1261 GAATTTCTGCTTCTCAGAGAAATAGACTTACTGCGCAGTATCTCATGAGCTTTA 1320  
1380 GAATTTCTGCTTCTCAGAGAAATAGACTTACTGCGCAGTATCTCATGAGCTTTA 1439  
1321 ATATGTAAAGTGAAGAGTTCCTCAATCAATGATGAGATTAATTTGAGACAAATA 1380  
1440 ATATGTAAAGTGAAGAGTTCCTCAATCAATGATGAGATTAATTTGAGACAAATA 1499  
1381 TTTGGGAAACCTATGAGGAAAGAGGCAAGCCTCCCACTTAAAGCATGTAACATAAT 1440  
1500 TTTGGGAAACCTATGAGGAAAGAGGCAAGCCTCCCACTTAAAGCATGTAACATAAT 1559  
1441 CTATTTATGAGGACATTTGTTACTGAGCAGACAGATTAATCAAGAGCCTCCCAAT 1500  
1560 CTATTTATGAGGACATTTGTTACTGAGCAGACAGATTAATCAAGAGCCTCCCAAT 1619  
1501 AAATTTAAAGCTTAAAGAGAGCTTACATCAGGCTTCACTGAGAGATTTTATCAAGAA 1560  
1620 AAATTTAAAGCTTAAAGAGAGCTTACATCAGGCTTCACTGAGAGATTTTATCAAGAA 1679  
1561 GCAGATTTGGCAGTTGCAAAAGACTCCTGAAATGATTAATCAGAGGAACTTAACCAAGGAG 1620  
1680 GCAGATTTGGCAGTTGCAAAAGACTCCTGAAATGATTAATCAGAGGAACTTAACCAAGGAG 1739  
1621 CAGAATGCTCAAGTATGATATTTACTAATAGTGTGATGAGATTAACCAAAAGGATGAT 1680  
1740 CAGAATGCTCAAGTATGATATTTACTAATAGTGTGATGAGATTAACCAAAAGGATGAT 1799  
1681 TCTATTCAAGATGAGAAAAATCTTAACCAATAGATCACTGAGAAAAAGATCTGCTTTC 1740  
1800 TCTATTCAAGATGAGAAAAATCTTAACCAATAGATCACTGAGAAAAAGATCTGCTTTC 1859  
1741 AAACGAAAGCTGAACCTTAAAGCAGCAGTATTAACCAATAGATCACTGAGAAAAAGATCTGCTTTC 1800  
1860 AAACGAAAGCTGAACCTTAAAGCAGCAGTATTAACCAATAGATCACTGAGAAAAAGATCTGCTTTC 1919  
1801 CACAATTTCAAAAGCACCTTAAAGAGATAGGCTGAGAGGAAAGCTTCTTACAGGAGATATT 1860  
1920 CACAATTTCAAAAGCACCTTAAAGAGATAGGCTGAGAGGAAAGCTTCTTACAGGAGATATT 1979  
1861 CATGCGCTTGAACCTAGTACGTAAGAAATTAACCCACCTAATTTGATGATGAGAA 1920  
1980 CATGCGCTTGAACCTAGTACGTAAGAAATTAACCCACCTAATTTGATGATGAGAA 2039  
1921 ATTGATAGTGTCTTACGATGAGATGAAGAAAAAGTAAACCAATAGGCGAGTC 1980  
2040 ATTGATAGTGTCTTACGATGAGATGAAGAAAAAGTAAACCAATAGGCGAGTC 2099



Db	4280	GACTCTCAGGGCATCTCTCTCAGAGTGACATTTTAAACCACTCAGCAGAGGATACATG	4313
Qy	4201	CAACATAACCTGATTAAGCTCCAGCAGAAATGCGTGAACATAAGAGCTGTCTTGAACAG	4260
Db	4320	CAACATAACCTGATTAAGCTCCAGCAGAAATGCGTGAACATAAGAGCTGTCTTGAACAG	4379
Qy	4261	CATGGAGACCGACGCTTCTAACAGCAACCTTCATCAATAGGAGCTCTTGCGCCCTTGAG	4320
Db	4380	CATGGAGACCGACGCTTCTAACAGCAACCTTCATCAATAGGAGCTCTTGCGCCCTTGAG	4439
Qy	4321	GACCTGCGAAATCCAGAACAAAGCACAATCAGAAAAAGCAGTATTAACTTCACAGAAAAAGT	4380
Db	4440	GACCTGCGAAATCCAGAACAAAGCACAATCAGAAAAAGCAGTATTAACTTCACAGAAAAAGT	4499
Qy	4381	AGTGAATACCTATTAGCCAGATTCAGAGAGCCCTTCTGCTGACAAAGTTTGAGGTGCT	4440
Db	4500	AGTGAATACCTATTAGCCAGATTCAGAGAGCCCTTCTGCTGACAAAGTTTGAGGTGCT	4559
Qy	4441	GCAGATAGTTCACCAAGTAAAAATTAACACAGAGGGAAGGTCATCCCTCTTAAA	4500
Db	4560	GCAGATAGTTCACCAAGTAAAAATTAACACAGAGGGAAGGTCATCCCTCTTAAA	4619
Qy	4501	TGCCCATCATAGATGATAGGTGGTACATGACAGTTGCTTGAGAGCTTCAGAAATAGA	4560
Db	4620	TGCCCATCATAGATGATAGGTGGTACATGACAGTTGCTTGAGAGCTTCAGAAATAGA	4679
Qy	4561	AACTAACCCATTCACAGAGAGCTCATTAAGTGTGATGAGAGCAACAGCTGGAA	4620
Db	4680	AACTAACCCATTCACAGAGAGCTCATTAAGTGTGATGAGAGCAACAGCTGGAA	4739
Qy	4621	GAGTCTGGGGCACAGATTTAGCGGAAACATCTTACTGSCCAAGGCAAGATCTGAGAGGA	4680
Db	4740	GAGTCTGGGGCACAGATTTAGCGGAAACATCTTACTGSCCAAGGCAAGATCTGAGAGGA	4799
Qy	4681	ACCCCTTAACCTGGAATCTGGAATCAGCCCTCTTCTGATGACCCCTGAACTGATCCTCTCT	4740
Db	4800	ACCCCTTAACCTGGAATCTGGAATCAGCCCTCTTCTGATGACCCCTGAACTGATCCTCTCT	4859
Qy	4741	GAAGACAGAGCCCCAGAGTCAAGCTGTGTGGCAACATACCATCTTCAACCTCGCATTTG	4800
Db	4860	GAAGACAGAGCCCCAGAGTCAAGCTGTGTGGCAACATACCATCTTCAACCTCGCATTTG	4919
Qy	4801	AAAGTCCCCAATGGAAGTTGCAAGATCTCCCAAGTCCAGCTGCTGCTCATATCACT	4860
Db	4920	AAAGTCCCCAATGGAAGTTGCAAGATCTCCCAAGTCCAGCTGCTGCTCATATCACT	4979
Qy	4861	GATACCTGCTGGGTATTAATGCATATGGAAGAAAGTGTGAGCAGGAGGAAGCCAGAAATTACA	4920
Db	4980	GATACCTGCTGGGTATTAATGCATATGGAAGAAAGTGTGAGCAGGAGGAAGCCAGAAATTACA	5039
Qy	4921	GCTTCAACAGAAAGGCTCAACAAAAAGATGTCATGTTGTTGTGGCTGACCCCGAAA	4980
Db	5040	GCTTCAACAGAAAGGCTCAACAAAAAGATGTCATGTTGTTGTGGCTGACCCCGAAA	5099
Qy	4981	GAATTTATGCTGCTGCTACAGTTTGGCCAGAAACACACATCACTTTAACTAATCTAATT	5040
Db	5100	GAATTTATGCTGCTGCTACAGTTTGGCCAGAAACACACATCACTTTAACTAATCTAATT	5159
Qy	5041	ACTGAAGAGACTACTCATGTTGTTATGAAAAACAGATGCTGATTTGTGTGGAACGGACA	5100
Db	5160	ACTGAAGAGACTACTCATGTTGTTATGAAAAACAGATGCTGATTTGTGTGGAACGGACA	5219
Qy	5101	CTGAATAATTTTCTAGGAATTTGCGGGAGGAAATGGGTAGTTAGCTATTCTTGCGGTACCC	5160
Db	5220	CTGAATAATTTTCTAGGAATTTGCGGGAGGAAATGGGTAGTTAGCTATTCTTGCGGTACCC	5279
Qy	5161	CAGTCTATTAAAGAAAGAAAAATCTGTAATGAGCATGATTTTGAAGTCAGAGAGATGTG	5220
Db	5380	CAGTCTATTAAAGAAAGAAAAATCTGTAATGAGCATGATTTTGAAGTCAGAGAGATGTG	5339
Qy	5221	GTCATATGGAAGAAACCAACAGGTCACAAAGCGACCAAGAGAAATCCCAAGGACAGAAATATC	5280
Db	5340	GTCATATGGAAGAAACCAACAGGTCACAAAGCGACCAAGAGAAATCCCAAGGACAGAAATATC	5399

Oy	5281	TTTCAGGGGCGCTGGAATCTGTTGCTATGAGGCCCTTCACCAACATGCCCCACAGATCAACTG	53440
Db	5400	TTTCAGGGGCGCTGGAATCTGTTGCTATGAGGCCCTTCACCAACATGCCCCACAGATCAACTG	54550
Oy	5341	GAATGATGATACACCTGTGTGCTCTCTCTGTGTGAAGAGACCTTTCATTCACCCCTT	54000
Db	5460	GAATGATGATGATACACCTGTGTGCTCTCTCTGTGTGAAGAGACCTTTCATTCACCCCTT	55190
Oy	5401	GGCAAGAGTGTCCACCCAAATTTGTTGTCGACCCAGATGCGCTGGACAGAGACAAATGGC	54660
Db	5520	GGCAAGAGTGTCCACCCAAATTTGTTGTCGACCCAGATGCGCTGGACAGAGACAAATGGC	55790
Oy	5461	TTCCATGCAATTTGGGCAAGATGTGTGAGCACACCTGTGTGTGACCCGAGAGATGGGTGTGGAC	55200
Db	5580	TTCCATGCAATTTGGGCAAGATGTGTGAGCACACCTGTGTGTGACCCGAGAGATGGGTGTGGAC	56339
Oy	5521	AGTGTAGCAGCTCTACCACTGTGCCAGAGAGCTGTGAGACCTACCTGATACCCAGATCCCCCAG	55800
Db	5640	AGTGTAGCAGCTCTACCACTGTGCCAGAGAGCTGTGAGACCTACCTGATACCCAGATCCCCCAG	56990
Oy	5581	AGCCACTAC 5589	
Db	5700	AGCCACTAC 5708	
RESULT 15			
AA117438			
ID	AA117438 standard; cDNA; 5914 BP.		
XX	AA117438;		
AC			
XX			
DT	30-SEP-1996 (first entry)		
XX			
DE	BRCA1 coding sequence.		
XX			
KW	Cancer therapy; breast and ovarian cancer predisposing gene; immunogen;		
KM	antibody production; germline alteration; probe; lesion neoplasia; human;		
KW	gene therapy; protein replacement therapy; protein mimetic; BRCA1; ds...		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	120..5711	
FT		/*tag= a	
FT		/product= BRCA1	
XX			
FN	W09605306-A2.		
XX			
PD	22-FEB-1996.		
XX			
PE	11-AUG-1995; 95WO-US10202.		
XX			
PR	07-JUN-1995;	95US-0483553.	
PR	12-AUG-1994;	94US-0289221.	
PR	02-SEP-1994;	94US-0300266.	
PR	16-SEP-1994;	94US-0308104.	
PR	29-NOV-1994;	94US-0348824.	
PR	24-MAR-1995;	95US-0409305.	
PR	07-JUN-1995;	95US-0480784.	
XX			
PA	(MYRI-) MYRIAD GENETICS INC.		
PA	(CANC-) CANCER INST.		
PA	(RECH-) CENT RECH DU CHUL.		
XX			
PI	Durocher F, Emi M, Nakamura Y, Simard J, Shattuck-Eidens DM;		
XX			
DR	WPI; 1996-139702/14.		
DR	P-PSDB; AAR81481.		
XX			
PT	New nucleic acid and polypeptide for mutant or polymorphic BRCA1		
PT	gene - for diagnosis and therapy of human breast and ovarian cancer		
PT	and for diagnosing pre-disposition to these cancers		

XX Claim 1; Page 119-128; 218bp; English.  
PS This sequence represents the cDNA of the human breast and ovarian cancer  
XX predisposing gene (BRCA1). Proteins encoded by mutations of this  
CC sequence (see AAT17433-117453 and AAT17455-117529) can be used as  
CC immunogens for antibody production. An additional BRCA1 mutant, with a T  
CC to G mutation and 59 nucleotide insert (encoding a protein which  
CC terminates at residue 75) is also claimed, however the exact location of  
CC the mutations is not specified so the sequence could not be created. The  
CC mutant BRCA1 genes have at least 1 mutation or polymorphism in comparison  
CC to this sequence. By detecting a germline alteration in this gene, a  
CC predisposition for breast and ovarian cancer can be diagnosed. In one  
CC method, BRCA1 mRNA isolated from a tissue sample from a subject has a  
CC probe, corresponding to a fragment of this sequence (or an  
CC allele-specific probe for a mutation of this sequence), added to it.  
CC The conditions allow for hybridisation of the probe to the mRNA, and any  
CC hybridisation which occurs is detected. Alternatively the BRCA1 gene in  
CC the tissue sample is isolated, and a shift in electrophoretic mobility  
CC of single stranded DNA from the sample on a non-denaturing  
CC polyacrylamide gel indicates a mutation. These methods of detection can  
CC also diagnose a lesion neoplasia associated with the BRCA1 locus. The  
CC methods may be used in gene therapy, protein replacement therapy and  
CC protein mimetics, and may be used to screen for drugs in cancer therapy.  
XX  
SQ Sequence 5914 BP; 2006 A; 1156 C; 1316 G; 1436 T; 0 other:  
Query Match 99.9%; Score 5585.8; DB 17; Length 5914;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 5587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 ATGGATTTATCTGCTCTTCCGCTTGAAGAAGTACAAATGCTATTATGTCAGAAA 60  
DB ATGGATTTATCTGCTCTTCCGCTTGAAGAAGTACAAATGCTATTATGTCAGAAA 179  
QY 61 ATCTTGAAGTGTCCATCTCTCTGAGTGTATCAAGAACCTGTCTCCACAAAGTGTGAC 120  
DB ATCTTGAAGTGTCCATCTCTCTGAGTGTATCAAGAACCTGTCTCCACAAAGTGTGAC 239  
QY 121 CACATATTTTGGCAATTTTGCATGCTGAAGCTCTCAACGAGAGAAAGGCTTTCACAG 180  
DB CACATATTTTGGCAATTTTGCATGCTGAAGCTCTCAACGAGAGAAAGGCTTTCACAG 299  
QY 240 CACATATTTTGGCAATTTTGCATGCTGAAGCTCTCAACGAGAGAAAGGCTTTCACAG 299  
DB CACATATTTTGGCAATTTTGCATGCTGAAGCTCTCAACGAGAGAAAGGCTTTCACAG 299  
QY 181 TGTCTTTATATGAAGATGATATACCAAAAGAGGCTTCAAGAAAGTACAGAGATTAGT 240  
DB TGTCTTTATATGAAGATGATATACCAAAAGAGGCTTCAAGAAAGTACAGAGATTAGT 359  
QY 300 TGTCTTTATATGAAGATGATATACCAAAAGAGGCTTCAAGAAAGTACAGAGATTAGT 359  
DB TGTCTTTATATGAAGATGATATACCAAAAGAGGCTTCAAGAAAGTACAGAGATTAGT 359  
QY 241 CAACTTGTGAGAGCTATTTGAATCATTTTGTCTTTCAGCTTGAACAGAGTTTGGAG 300  
DB CAACTTGTGAGAGCTATTTGAATCATTTTGTCTTTCAGCTTGAACAGAGTTTGGAG 419  
QY 301 TATGCAAAACAGCTATATTTTGCAGAAAAGAAATTAACCTCCCTCAACATCTAAAGAT 360  
DB TATGCAAAACAGCTATATTTTGCAGAAAAGAAATTAACCTCCCTCAACATCTAAAGAT 479  
QY 420 TATGCAAAACAGCTATATTTTGCAGAAAAGAAATTAACCTCCCTCAACATCTAAAGAT 479  
DB TATGCAAAACAGCTATATTTTGCAGAAAAGAAATTAACCTCCCTCAACATCTAAAGAT 479  
QY 361 GAAGTTTCTATCATCCAAAGATAGGCTACAGAAACCGTCCAAAGAGCTTCTACAGAGT 420  
DB GAAGTTTCTATCATCCAAAGATAGGCTACAGAAACCGTCCAAAGAGCTTCTACAGAGT 539  
QY 480 GAAGTTTCTATCATCCAAAGATAGGCTACAGAAACCGTCCAAAGAGCTTCTACAGAGT 539  
DB GAAGTTTCTATCATCCAAAGATAGGCTACAGAAACCGTCCAAAGAGCTTCTACAGAGT 539  
QY 421 GAACCCGAAATATCTTCTTTCAGAGAAACAGCTCTAGTGTCCAACTCTCTAACTTGGG 480  
DB GAACCCGAAATATCTTCTTTCAGAGAAACAGCTCTAGTGTCCAACTCTCTAACTTGGG 599  
QY 540 GAACCCGAAATATCTTCTTTCAGAGAAACAGCTCTAGTGTCCAACTCTCTAACTTGGG 599  
DB GAACCCGAAATATCTTCTTTCAGAGAAACAGCTCTAGTGTCCAACTCTCTAACTTGGG 599  
QY 481 ACTGTGAGAACCTGAGGAGCAAGAGCGGATCAACCTCAAAAGAGCTGTCTACAT 540  
DB ACTGTGAGAACCTGAGGAGCAAGAGCGGATCAACCTCAAAAGAGCTGTCTACAT 659  
QY 600 ACTGTGAGAACCTGAGGAGCAAGAGCGGATCAACCTCAAAAGAGCTGTCTACAT 659  
DB ACTGTGAGAACCTGAGGAGCAAGAGCGGATCAACCTCAAAAGAGCTGTCTACAT 659  
QY 541 GAATGGAGTCTGATCTTCTGAAGATACGTTAATAGGCAACTATTGCAAGTGGG 600  
DB GAATGGAGTCTGATCTTCTGAAGATACGTTAATAGGCAACTATTGCAAGTGGG 719  
QY 660 GAATGGAGTCTGATCTTCTGAAGATACGTTAATAGGCAACTATTGCAAGTGGG 719  
DB GAATGGAGTCTGATCTTCTGAAGATACGTTAATAGGCAACTATTGCAAGTGGG 719  
QY 601 GATCAAGATTTGTACAAATCACCCCTCAAGAGACGAGGATGAATCAGTTTGGATTCT 660  
DB GATCAAGATTTGTACAAATCACCCCTCAAGAGACGAGGATGAATCAGTTTGGATTCT 660

DB 720 GATCAAGATTTGTACAAATCACCCCTCAAGAGACGAGGATGAATCAGTTTGGATTCT 779  
QY 661 GCAGAAAAGAGCTCTTGTGATTTTCTGACAGCATGTAACTAATCAATCATCA 720  
DB GCAGAAAAGAGCTCTTGTGATTTTCTGACAGCATGTAACTAATCAATCATCA 839  
QY 721 CCCAGTAATATGATTTTGAACACACCATGAGAACGCTGACAGCTGAGAGCAAGAA 780  
DB CCCAGTAATATGATTTTGAACACACCATGAGAACGCTGACAGCTGAGAGCAAGAA 899  
QY 840 CCCAGTAATATGATTTTGAACACACCATGAGAACGCTGACAGCTGAGAGCAAGAA 899  
DB CCCAGTAATATGATTTTGAACACACCATGAGAACGCTGACAGCTGAGAGCAAGAA 899  
QY 781 TATCAGAGTATGTTCTTTTCAACTTGCATGTGAGACCATGTGGCAAAATCTATGCC 840  
DB TATCAGAGTATGTTCTTTTCAACTTGCATGTGAGACCATGTGGCAAAATCTATGCC 959  
QY 900 TATCAGAGTATGTTCTTTTCAACTTGCATGTGAGACCATGTGGCAAAATCTATGCC 959  
DB TATCAGAGTATGTTCTTTTCAACTTGCATGTGAGACCATGTGGCAAAATCTATGCC 959  
QY 841 AGCTCATTTACAGATGAGAGACAGCATTTATTAACCTCAAAAGAGCAAGATGATGAGA 900  
DB AGCTCATTTACAGATGAGAGACAGCATTTATTAACCTCAAAAGAGCAAGATGATGAGA 1019  
QY 901 AAGGCTGAATTTCTGTAATAAAGCAACAGCGCTGTGAGCAAGAGCCACATTAACAGA 960  
DB AAGGCTGAATTTCTGTAATAAAGCAACAGCGCTGTGAGCAAGAGCCACATTAACAGA 1079  
QY 1020 AAGGCTGAATTTCTGTAATAAAGCAACAGCGCTGTGAGCAAGAGCCACATTAACAGA 1079  
DB AAGGCTGAATTTCTGTAATAAAGCAACAGCGCTGTGAGCAAGAGCCACATTAACAGA 1079  
QY 961 TGGGCTGGAAGTAAAGCAACATGTAATGATGAGCGGACATCCAGCACAGAAAAAGTAA 1020  
DB TGGGCTGGAAGTAAAGCAACATGTAATGATGAGCGGACATCCAGCACAGAAAAAGTAA 1139  
QY 1080 TGGGCTGGAAGTAAAGCAACATGTAATGATGAGCGGACATCCAGCACAGAAAAAGTAA 1139  
DB TGGGCTGGAAGTAAAGCAACATGTAATGATGAGCGGACATCCAGCACAGAAAAAGTAA 1139  
QY 1021 GATCTGAATGCTGATCCCTGTGTGAGAGAAAAAGTAAATGATGAGCAAACTGCCATGC 1080  
DB GATCTGAATGCTGATCCCTGTGTGAGAGAAAAAGTAAATGATGAGCAAACTGCCATGC 1199  
QY 1140 GATCTGAATGCTGATCCCTGTGTGAGAGAAAAAGTAAATGATGAGCAAACTGCCATGC 1199  
DB GATCTGAATGCTGATCCCTGTGTGAGAGAAAAAGTAAATGATGAGCAAACTGCCATGC 1199  
QY 1081 TCAGAGAAATCTAGAGATCTAGAGATGTTCTTGTGATTAACACTTAATAGCAGATTTCAG 1140  
DB TCAGAGAAATCTAGAGATCTAGAGATGTTCTTGTGATTAACACTTAATAGCAGATTTCAG 1259  
QY 1200 TCAGAGAAATCTAGAGATCTAGAGATGTTCTTGTGATTAACACTTAATAGCAGATTTCAG 1259  
DB TCAGAGAAATCTAGAGATCTAGAGATGTTCTTGTGATTAACACTTAATAGCAGATTTCAG 1259  
QY 1141 AAGGTTAATGAGGTTTTCAGAAATGATGAACCTGTATGATCTATCAATCATGAT 1200  
DB AAGGTTAATGAGGTTTTCAGAAATGATGAACCTGTATGATCTATCAATCATGAT 1319  
QY 1260 AAGGTTAATGAGGTTTTCAGAAATGATGAACCTGTATGATCTATCAATCATGAT 1319  
DB AAGGTTAATGAGGTTTTCAGAAATGATGAACCTGTATGATCTATCAATCATGAT 1319  
QY 1201 GGGGAGTCTGAATCAATGAGCAAGTAGCTGATGATTTGAGAGCTGTTAATGAGTAGAT 1260  
DB GGGGAGTCTGAATCAATGAGCAAGTAGCTGATGATTTGAGAGCTGTTAATGAGTAGAT 1379  
QY 1320 GGGGAGTCTGAATCAATGAGCAAGTAGCTGATGATTTGAGAGCTGTTAATGAGTAGAT 1379  
DB GGGGAGTCTGAATCAATGAGCAAGTAGCTGATGATTTGAGAGCTGTTAATGAGTAGAT 1379  
QY 1261 GAATATTTCTGTTCTTCAAGAGAAATATAGACTTACTGCGCAGTGTATCATGAGGCTTTA 1320  
DB GAATATTTCTGTTCTTCAAGAGAAATATAGACTTACTGCGCAGTGTATCATGAGGCTTTA 1439  
QY 1380 GAATATTTCTGTTCTTCAAGAGAAATATAGACTTACTGCGCAGTGTATCATGAGGCTTTA 1439  
DB GAATATTTCTGTTCTTCAAGAGAAATATAGACTTACTGCGCAGTGTATCATGAGGCTTTA 1439  
QY 1321 ATATGTAAGAGTGAAGAGTTTCACTCCAAATCAGTGAAGATATATTTGAGCAAAATTA 1380  
DB ATATGTAAGAGTGAAGAGTTTCACTCCAAATCAGTGAAGATATATTTGAGCAAAATTA 1499  
QY 1440 ATATGTAAGAGTGAAGAGTTTCACTCCAAATCAGTGAAGATATATTTGAGCAAAATTA 1499  
DB ATATGTAAGAGTGAAGAGTTTCACTCCAAATCAGTGAAGATATATTTGAGCAAAATTA 1499  
QY 1381 TTTGGGAAAACTATGAGGAAAGGCAAGGCTCCCACTTAAAGCATGTAACCTGAAT 1440  
DB TTTGGGAAAACTATGAGGAAAGGCAAGGCTCCCACTTAAAGCATGTAACCTGAAT 1559  
QY 1500 TTTGGGAAAACTATGAGGAAAGGCAAGGCTCCCACTTAAAGCATGTAACCTGAAT 1559  
DB TTTGGGAAAACTATGAGGAAAGGCAAGGCTCCCACTTAAAGCATGTAACCTGAAT 1559  
QY 1441 CTATATTAAGAGACATTTTCTAGAGCCACAGATATTAACAAGAGGCTCCCTCACAAAT 1500  
DB CTATATTAAGAGACATTTTCTAGAGCCACAGATATTAACAAGAGGCTCCCTCACAAAT 1619  
QY 1560 CTATATTAAGAGACATTTTCTAGAGCCACAGATATTAACAAGAGGCTCCCTCACAAAT 1619  
DB CTATATTAAGAGACATTTTCTAGAGCCACAGATATTAACAAGAGGCTCCCTCACAAAT 1619  
QY 1501 AAATTAAGAGCTGAAGAGAGCTACATCAGGCTTCTATCTGAGAGATTTTATCAAGAAA 1560  
DB AAATTAAGAGCTGAAGAGAGCTACATCAGGCTTCTATCTGAGAGATTTTATCAAGAAA 1679  
QY 1620 AAATTAAGAGCTGAAGAGAGCTACATCAGGCTTCTATCTGAGAGATTTTATCAAGAAA 1679  
DB AAATTAAGAGCTGAAGAGAGCTACATCAGGCTTCTATCTGAGAGATTTTATCAAGAAA 1679  
QY 1561 GCAGATTTGGCAGTTGCAAAAAGACTCTGAAATGATTAATCAAGGAGCTTAACCAAGGAG 1620  
DB GCAGATTTGGCAGTTGCAAAAAGACTCTGAAATGATTAATCAAGGAGCTTAACCAAGGAG 1739  
QY 1680 GCAGATTTGGCAGTTGCAAAAAGACTCTGAAATGATTAATCAAGGAGCTTAACCAAGGAG 1739  
DB GCAGATTTGGCAGTTGCAAAAAGACTCTGAAATGATTAATCAAGGAGCTTAACCAAGGAG 1739  
QY 1621 CAGAAATGTCAGAGTGAATATTAATAATAGTGTCTATGAGAAATTAACCAAGGCTGAT 1680  
DB CAGAAATGTCAGAGTGAATATTAATAATAGTGTCTATGAGAAATTAACCAAGGCTGAT 1799  
QY 1740 CAGAAATGTCAGAGTGAATATTAATAATAGTGTCTATGAGAAATTAACCAAGGCTGAT 1799  
DB CAGAAATGTCAGAGTGAATATTAATAATAGTGTCTATGAGAAATTAACCAAGGCTGAT 1799  
QY 1681 TCTATTTCAAGATGAGAAAAATCTTAACCAATTAAGATCTCTGAGAAATCTGCTTTC 1740  
DB TCTATTTCAAGATGAGAAAAATCTTAACCAATTAAGATCTCTGAGAAATCTGCTTTC 1859

1741 AAAACGAAAGCTGAACTATTAAGCAGAGTATTAAGCAATATGGAACCTGCAATTAAATATC 1800  
1860 AAAAGCAAGAGCTGAACCTATTAAGCAGAGTATTAAGCAATATGGAACCTGCAATTAAATATC 1919  
1801 CACAAATTCAAAGACCTTAAGAAATAGAGCTGAGAGAAAGTCTTACAGAGCATATT 1860  
1920 CACAATTCAAAGACCTTAAGAAATAGAGCTGAGAGAAAGTCTTACAGAGCATATT 1979  
1861 CATGGCTTGAAGCTAGTACATTAAGAAATCTAAGCCCACTAATGCTACTGAAATTCGA 1920  
1980 CATGGCTTGAAGCTAGTACATTAAGAAATCTAAGCCCACTAATGCTACTGAAATTCGA 2039  
1921 AATGATAGTGTGTTCTAGCAGTGAAGATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 1980  
2040 AATGATAGTGTGTTCTAGCAGTGAAGATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 2099  
1981 AGGCACAGCAGAAACCTACAACTCATGGAAGGTAAGAAACCTCAACTGAGCCAGAAAG 2040  
2100 AGGCACAGCAGAAACCTACAACTCATGGAAGGTAAGAAACCTCAACTGAGCCAGAAAG 2159  
2041 AGTAACAGCCAAATGAACAGACAGTAAGAAAGACATGACAGTACTTCCAGAGCTG 2100  
2160 AGTAACAGCCAAATGAACAGACAGTAAGAAAGACATGACAGTACTTCCAGAGCTG 2219  
2101 AAGTTAAACAATGCACTGTTCTTTTACTAAGTGTTCAAATACAGTGAACCTTAAGAA 2160  
2220 AAGTTAAACAATGCACTGTTCTTTTACTAAGTGTTCAAATACAGTGAACCTTAAGAA 2279  
2161 TTTGTCATCTCTAAGCTTCCAGAGAGAAAGAAAGAAAGAAAGAAAGAAAGTGAAGTG 2220  
2280 TTTGTCATCTCTAAGCTTCCAGAGAGAAAGAAAGAAAGAAAGAAAGTGAAGTG 2339  
2221 TCTAATTAATGCTGAAGACCCCAAGAGATCTCATGTTAAGTGAGAGAAAGGTTTGAACCT 2280  
2340 TCTAATTAATGCTGAAGACCCCAAGAGATCTCATGTTAAGTGAGAGAAAGGTTTGAACCT 2399  
2281 GAAAGATCTGTAGAGAGTGAAGATTAATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 2340  
2400 GAAAGATCTGTAGAGAGTGAAGATTAATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 2459  
2341 GAAAGATCTGTAGAGAGTGAAGATTAATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 2400  
2460 GAAAGATCTGTAGAGAGTGAAGATTAATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 2519  
2401 TGTGTGAGTCACTGTGCAACATTTGAAACCCCAAGAGACTAATTCATGCTGTTTCCAA 2460  
2520 TGTGTGAGTCACTGTGCAACATTTGAAACCCCAAGAGACTAATTCATGCTGTTTCCAA 2579  
2461 GATTAATGAAATGACACAGAGGCTTTAAGTATCCATTGGGACATGAACTTAACACAGT 2520  
2580 GATTAATGAAATGACACAGAGGCTTTAAGTATCCATTGGGACATGAACTTAACACAGT 2639  
2521 CGGGAACAGAGATGAAATGGAAGAAAGTGAACCTGATGCTCAGATTTGCAAGATTA 2580  
2640 CGGGAACAGAGATGAAATGGAAGAAAGTGAACCTGATGCTCAGATTTGCAAGATTA 2699  
2581 TTCAAGTGTTCAAAGCCGAGTCAATTTGCTGTTTCAATTCAGAGAAATGAGAGAG 2640  
2700 TTCAAGTGTTCAAAGCCGAGTCAATTTGCTGTTTCAATTCAGAGAAATGAGAGAG 2759  
2641 GAATGTGCAACATTTCTGCCCACCTCTGGGCTTTAAAGAAACAAAGTCCAAAGCTCACT 2700  
2760 GAATGTGCAACATTTCTGCCCACCTCTGGGCTTTAAAGAAACAAAGTCCAAAGCTCACT 2819  
2701 TTTGATGTGAACAAAGAGAGAAATCAAGAGAAAGATGAGTCTAATATCAAGCTGTA 2760  
2820 TTTGATGTGAACAAAGAGAGAAATCAAGAGAAAGATGAGTCTAATATCAAGCTGTA 2879  
2761 CAGACAGTTAATATGACTGAGAGGCTTCTGTTGTTGTCAGAAAGATTAAGCAGTGTAT 2820  
2880 CAGACAGTTAATATGACTGAGAGGCTTCTGTTGTTGTCAGAAAGATTAAGCAGTGTAT 2939

2821 AATGCCAAATGTAGTATCAAGAGAGGCTCTAGGTTTGTCTATCATCTCAGTTCAGAGGC 2880  
2940 AATGCCAAATGTAGTATCAAGAGAGGCTCTAGGTTTGTCTATCATCTCAGTTCAGAGGC 2999  
2881 AACGAACCTGAGACTATTAATCCAAATTAACATGAGACTTTTACAAACCCATATCTGTA 2940  
3000 AACGAACCTGAGACTATTAATCCAAATTAACATGAGACTTTTACAAACCCATATCTGTA 3059  
2941 CCACCACTTTTCCCTCAGTCAATTTGTTAAACTAAATGTAAGAAATCTGCTAGAG 3000  
3060 CCACCACTTTTCCCTCAGTCAATTTGTTAAACTAAATGTAAGAAATCTGCTAGAG 3119  
3001 GAAACCTTGAGAACATTAATGTCACCTGGAAGAAATGGAATGAGAACATTTCA 3060  
3120 GAAACCTTGAGAACATTAATGTCACCTGGAAGAAATGGAATGAGAACATTTCA 3179  
3061 AGTACAGTGAACCAATTAAGCGTAAATTAACATTAAGAAATGTTTAAAGAGCCAGC 3120  
3180 AGTACAGTGAACCAATTAAGCGTAAATTAACATTAAGAAATGTTTAAAGAGCCAGC 3239  
3121 TCAGCAATTAATGAAGTGAAGTTCAGTACTAATGAAGTGGGCTCCAGATTAATGA 3180  
3240 TCAGCAATTAATGAAGTGAAGTTCAGTACTAATGAAGTGGGCTCCAGATTAATGA 3299  
3181 ATAGGTTCCAGTATGAAGAAACATTAAGCAGAGAACTAGTAGAAACAGAGGCCAAATTTG 3240  
3300 ATAGGTTCCAGTATGAAGAAACATTAAGCAGAGAACTAGTAGAAACAGAGGCCAAATTTG 3359  
3241 AATGCTATCTGATGATTAAGGCTTTTGAACCTGAGTCTAATTAACAAAGTCTTCTGGA 3300  
3360 AATGCTATCTGATGATTAAGGCTTTTGAACCTGAGTCTAATTAACAAAGTCTTCTGGA 3419  
3301 AGTAATGTAACATCTGTAATTAAGAAAGAAAGTGAAGAAAGTGAAGTGAAGTGAAGTGAAGT 3360  
3420 AGTAATGTAACATCTGTAATTAAGAAAGAAAGTGAAGAAAGTGAAGTGAAGTGAAGTGAAGT 3479  
3361 AATACAGATTTCTCCATATCTGATTTCAATTAAGTGAAGAAAGCTATGGAAGTGAAGT 3420  
3480 AATACAGATTTCTCCATATCTGATTTCAATTAAGTGAAGAAAGCTATGGAAGTGAAGT 3539  
3421 CATGATCTCAGGTTTGTCTGAGACACCTGATGACCTGTTAATGATGATGATGATGATGATGAT 3480  
3540 CATGATCTCAGGTTTGTCTGAGACACCTGATGACCTGTTAATGATGATGATGATGATGATGAT 3599  
3481 GAAAGATCTAGTGTGCTGAATGAATGACATTAAGAAAGTCTGCTGTTTGAAGAAAGC 3540  
3600 GAAAGATCTAGTGTGCTGAATGAATGACATTAAGAAAGTCTGCTGTTTGAAGAAAGC 3659  
3541 GTCCAGAAAGGAGACCTTAGAGAGAGTCTAGAGCTTACCCATTAACATTTGGCTCAG 3600  
3660 GTCCAGAAAGGAGACCTTAGAGAGAGTCTAGAGCTTACCCATTAACATTTGGCTCAG 3719  
3601 GGTAAACCGAAGAGGGGCCAAGAAATTAAGATTCAGAGAGAACTTAATGAGAGAT 3660  
3720 GGTAAACCGAAGAGGGGCCAAGAAATTAAGATTCAGAGAGAACTTAATGAGAGAT 3779  
3661 GAAGAGCTTCCCTGCTTCCACACATTTGATTTGGTAAAGTAAATTAATTAATTAATTAATTAAT 3720  
3780 GAAGAGCTTCCCTGCTTCCACACATTTGATTTGGTAAAGTAAATTAATTAATTAATTAATTAAT 3839  
3721 TCTACTAGGCAATGACACCGTGTCTACGAGTGTCTGTCTAAGAACACAGAGAGAAATTTA 3780  
3840 TCTACTAGGCAATGACACCGTGTCTACGAGTGTCTGTCTAAGAACACAGAGAGAAATTTA 3899  
3781 TTATCATTTGAAGATAGCTTTAATGACTGCAATTAACAGATTAATTTGGCAAGAGCATCT 3840  
3900 TTATCATTTGAAGATAGCTTTAATGACTGCAATTAACAGATTAATTTGGCAAGAGCATCT 3959  
3841 CAGGAACATCACTAGTGAAGAAACAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3900  
3960 CAGGAACATCACTAGTGAAGAAACAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4019  
3901 AGTGAATGGAAGACTGCTGCAATTAACAAACACCCAGAGATCCTTCTTGATTTGGTTCT 3960









GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 26, 2003, 16:19:58 ; Search time 7050.55 Seconds  
(without alignments)  
12838.232 Million cell updates/sec

Title: US-09-734-672-3\_COPY\_120\_5708

Perfect score: 5589  
Sequence: 1 ATGAGATTATCTGCTCTCTG.....AGATCCCAACAGACACTAC 5589

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

```

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estln:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1960.8	35.1	2101	11	BC012577	Homo sapi
2	1955	35.0	2090	11	BC030969	Homo sapi
3	861.4	15.4	962	14	B0068830	AGENCOURT
4	843	15.1	1089	13	BM452288	AGENCOURT
5	811.8	14.5	845	9	AU122476	AU122476
6	761.8	13.6	783	9	AU142729	AU142729

7	740.4	13.2	743	12	BG681276
8	688	12.3	702	9	AU125312
9	686.6	12.3	739	12	BF508987
10	686.4	12.3	987	14	BM800251
11	666.6	11.9	899	14	BQ422380
12	657	11.8	900	12	BF791668
13	619.2	11.1	921	12	BG178466
14	616.2	11.0	878	14	BQ215100
15	603.8	10.8	630	9	A1992040
16	592.6	10.6	800	12	BF794879
17	591.4	10.6	627	9	A1589028
18	591	10.6	724	9	A1915085
19	590	10.6	906	14	BQ878445
20	590	10.6	909	14	BQ676829
21	590	10.6	933	14	BQ683955
22	587	10.5	747	12	BG777447
23	583	10.4	947	14	BQ679749
24	578	10.3	910	14	BQ677666
25	575	10.3	602	10	BE043993
26	561.6	10.0	938	12	BG257190
27	560.4	10.0	675	9	AL704228
28	545.2	9.8	638	13	BM042282
29	536.8	9.6	638	10	BE264293
30	531.6	9.5	563	10	AW295197
31	530.2	9.5	585	10	AW968546
32	518	9.3	666	10	AW968720
33	499.8	8.9	509	10	AW504244
34	492.8	8.8	518	9	AA804632
35	491.8	8.8	523	9	AA702344
36	491.2	8.8	509	10	BE018878
37	479.4	8.6	1220	14	BQ214737
38	472.2	8.4	487	9	AA812019
39	450.6	8.1	575	10	BE564528
40	449.8	8.0	453	9	A1684595
41	444	7.9	444	9	AA486004
42	437.4	7.8	488	9	AA814998
43	421.8	7.5	429	9	AA484941
44	420.6	7.5	459	9	A1040685
45	418.2	7.5	443	14	BQ308670

## ALIGNMENTS

RESULT 1  
BC012577  
LOCUS 2101 bp mRNA linear HTC 20-AUG-2001  
DEFINITION Homo sapiens, clone IMAGE:399658, mRNA.  
ACCESSION BC012577  
VERSION BC012577.1 GI:15214876  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 2101)  
AUTHORS Strausberg, R.  
TITLE Direct Submission  
JOURNAL Submitted (15-AUG-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
REMARK  
COMMENT NTH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: CLONTECH Laboratories, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNU)  
DNA Sequencing by: Institute for Systems Biology  
<http://www.systemsbio.org>  
contact: [amadansystemsbiology.org](mailto:amadansystemsbiology.org)  
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia Greene, Mark Keltman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRRL Plate: 19 Row: a Column: 18  
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6552298  
 This clone has the following problem: incomplete processing.

## FEATURES

source

1.. 2101  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:399658"  
 /tissue\_type="Bladder, carcinoma"  
 /clone\_lib="NIH\_MGC\_53"  
 /lab\_host="DH10B"  
 /note="Vector: pDNR-LIB"

BASE COUNT 756 a 385 c 452 g 508 t  
 ORIGIN

Query Match 35.1% Score 1960.8; DB 11; Length 2101;  
 Best Local Similarity 99.6%; Pred. No. 0;  
 Matches 1965; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

1 ATGATTTATCTGCTCTTCGCGTTGAAGAAGTACAAATGTCAATTAATGTCAGAAA 60  
 |||||  
 124 ATGATTTATCTGCTCTTCGCGTTGAAGAAGTACAAATGTCAATTAATGTCAGAAA 183  
 |||||  
 61 ATCTAGAGTCTCCATCTGCTGAGTATCAAGAACCTGCTCCCAAAAGTGGAC 120  
 |||||  
 184 ATCTAGAGTCTCCATCTGCTGAGTATCAAGAACCTGCTCCCAAAAGTGGAC 243  
 |||||  
 121 CACATATTTTGAATTTTGCATGCTGGAACCTTCTACAGAAAGAGGCGCTTCACAG 180  
 |||||  
 244 CACATATTTTGAATTTTGCATGCTGGAACCTTCTACAGAAAGAGGCGCTTCACAG 303  
 |||||  
 181 TGTCTTTATGTAAGATATATTAACCAAAAGAGGCTTACAGAAAGTACAGATTAGT 240  
 |||||  
 304 TGTCTTTATGTAAGATATATTAACCAAAAGAGGCTTACAGAAAGTACAGATTAGT 363  
 |||||  
 241 CAACCTTTGAGAGCTATTTGAATTCATTTGCTTTCAGCTTACAGAGTTTGGAG 300  
 |||||  
 364 CAACCTTTGAGAGCTATTTGAATTCATTTGCTTTCAGCTTACAGAGTTTGGAG 423  
 |||||  
 301 TATGCAAAACAGCTATTAATTTTGCACAAAAGAAATTAAGTCTTCTGACATCTAAAGAT 360  
 |||||  
 424 TATGCAAAACAGCTATTAATTTTGCACAAAAGAAATTAAGTCTTCTGACATCTAAAGAT 483  
 |||||  
 361 GAACTTTCTATCATCCAAAGTATGGGCTTACAGAAACCGTCCAAAAGACTTCTACAGAGT 420  
 |||||  
 484 GAACTTTCTATCATCCAAAGTATGGGCTTACAGAAACCGTCCAAAAGACTTCTACAGAGT 543  
 |||||  
 421 GAACCCGAAAATCTCTCTTGCAGAAAACAGCTCTGATGTCACACTCTTAACCTTTGG 480  
 |||||  
 544 GAACCCGAAAATCTCTCTTGCAGAAAACAGCTCTGATGTCACACTCTTAACCTTTGG 603  
 |||||  
 481 ACTGTGAGAACTCTGAGGCAAAAGCAGCGATACAACTCAAAAGACGCTGTCTACATT 540  
 |||||  
 604 ACTGTGAGAACTCTGAGGCAAAAGCAGCGATACAACTCAAAAGACGCTGTCTACATT 663  
 |||||  
 541 GAATTGGATCTGATTTCTTCTGAAGTACCGTTAATTAAGCAACTTATTGCACTGGGA 600  
 |||||  
 664 GAATTGGATCTGATTTCTTCTGAAGTACCGTTAATTAAGCAACTTATTGCACTGGGA 723  
 |||||  
 601 GATCAAGATTTGTTCAAAATCACCCCTCAAGGAAACAGGATGAATCACTGTTGATTC 660  
 |||||  
 724 GATCAAGATTTGTTCAAAATCACCCCTCAAGGAAACAGGATGAATCACTGTTGATTC 783  
 |||||  
 661 GCAAAAAAGGCTGCTTGTGAATTTTCTGAGAGGATGTAAACAAATCTACTGAATCATCA 720  
 |||||  
 784 GCAAAAAAGGCTGCTTGTGAATTTTCTGAGAGGATGTAAACAAATCTACTGAATCATCA 843  
 |||||  
 721 CCCAGTAAATGATTTTGAACACCACTGGAAGCGTGCAGCTGAGAGGCATCCAGAAAG 780  
 |||||

844 CCCAGTAAATGATTTTGAACACCACTGGAAGCGTGCAGCTGAGAGGCATCCAGAAAAG 903  
 |||||  
 781 TATCAGAGTATGCTCTTCAAACTGCTATGAGGCAATGGGCAACAAATCATGCC 840  
 |||||  
 904 TATCAGAGTATGCTCTTCAAACTGCTATGAGGCAATGGGCAACAAATCATGCC 963  
 |||||  
 841 AGCTCATTTACAGATGAGAACAGCAGTTTATTACTACTAAAGACAGATGATGTAGAA 900  
 |||||  
 964 AGCTCATTTACAGATGAGAACAGCAGTTTATTACTACTAAAGACAGATGATGTAGAA 1023  
 |||||  
 901 AAGGCTGAATTTCTGATTAATAAGCAAAACCGCTGCTTACAGAGACCAATACAGA 960  
 |||||  
 1024 AAGGCTGAATTTCTGATTAATAAGCAAAACCGCTGCTTACAGAGACCAATACAGA 1083  
 |||||  
 961 TGGGCTGGAAGTAAAGAAACATGTAATGATAGCGGACCTCCAGCACGAAAAAGGTA 1020  
 |||||  
 1084 TGGGCTGGAAGTAAAGAAACATGTAATGATAGCGGACCTCCAGCACGAAAAAGGTA 1143  
 |||||  
 1021 GATCTGAATGCTGATCCCTGCTGAGAGAAAGAAATGAATTAACAGAACTGCCATGC 1080  
 |||||  
 1144 GATCTGAATGCTGATCCCTGCTGAGAGAAAGAAATGAATTAACAGAACTGCCATGC 1203  
 |||||  
 1081 TCAGAGAAATCTAGAGATCTGAGATGTTCTTGGATTAACCTAAATAGCAGATTGAG 1140  
 |||||  
 1204 TCAGAGAAATCTAGAGATCTGAGATGTTCTTGGATTAACCTAAATAGCAGATTGAG 1263  
 |||||  
 1141 AAGGTAATGAGAGGTTTCCAGAAAGTATGTAAGTATGTTGATGATGATCATGAT 1200  
 |||||  
 1264 AAGGTAATGAGAGGTTTCCAGAAAGTATGTAAGTATGTTGATGATGATCATGAT 1323  
 |||||  
 1201 GGGGAGTCTGAATCAAAATCCCAAGTATGATTTGGAGCTTCTAATGAGTAGAT 1260  
 |||||  
 1324 GGGGAGTCTGAATCAAAATCCCAAGTATGATTTGGAGCTTCTAATGAGTAGAT 1383  
 |||||  
 1261 GAATATTTGCTTCTTCAAGAAATTAAGACTTACAGGACAGTATGATGAGGCTTTA 1320  
 |||||  
 1384 GAATATTTGCTTCTTCAAGAAATTAAGACTTACAGGACAGTATGATGAGGCTTTA 1443  
 |||||  
 1321 ATATGTAAGAGTAAGAGTCTACTCCAAATAGTAGAGATTAATTAAGACAAATA 1380  
 |||||  
 1444 ATATGTAAGAGTAAGAGTCTACTCCAAATAGTAGAGATTAATTAAGACAAATA 1503  
 |||||  
 1381 TTTGGGAAAACCTATTCGGAAGAGGCAAGCTCCCAACTTAAGCATGTAAAT 1440  
 |||||  
 1504 TTTGGGAAAACCTATTCGGAAGAGGCAAGCTCCCAACTTAAGCATGTAAAT 1563  
 |||||  
 1441 CTATATATGAGGACATTTGCTACTGAGCCAGCAATTAACAGAGCGTCCCTGCAAT 1500  
 |||||  
 1564 CTATATATGAGGACATTTGCTACTGAGCCAGCAATTAACAGAGCGTCCCTGCAAT 1623  
 |||||  
 1501 AATTTAAAGCGTAAAGAGACCTACATAGGCTTCACTGAGAGTTTATCAAGAAA 1560  
 |||||  
 1624 AATTTAAAGCGTAAAGAGACCTACATAGGCTTCACTGAGAGTTTATCAAGAAA 1683  
 |||||  
 1561 GCAGATTTGGCACTTCAAAAAGACTCTGAATGATTAATCAGGAACTAACCAACGGAG 1620  
 |||||  
 1684 GCAGATTTGGCACTTCAAAAAGACTCTGAATGATTAATCAGGAACTAACCAACGGAG 1743  
 |||||  
 1621 CAGATTTGCAAGTGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1680  
 |||||  
 1744 CAGATTTGCAAGTGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1803  
 |||||  
 1681 TCTATTCAGAAATGAGAAATATCTTAACCAATAGAAATCACTGAGAAAGAAATCTGCTTC 1740  
 |||||  
 1804 TCTATTCAGAAATGAGAAATATCTTAACCAATAGAAATCACTGAGAAAGAAATCTGCTTC 1863  
 |||||  
 1741 AAAAGCAAGCTGAACCTATAAGCAGAGTAAACCAATAGAAATCACTGAGAAATCAATC 1800  
 |||||  
 1864 AAAAGCAAGCTGAACCTATAAGCAGAGTAAACCAATAGAAATCACTGAGAAATCAATC 1923  
 |||||  
 1801 CACAAATTTCAAAAGCACTTAAGAAATAGGCTGAGAGGAAAGCTTTCTACAGGCATATT 1860  
 |||||  
 1924 CACAAATTTCAAAAGCACTTAAGAAATAGGCTGAGAGGAAAGCTTTCTACAGGCATATT 1983  
 |||||

QY	1861	CATCGCGTTGAACTAGAGTACAGTACAGTAAACCTAGCCCACTATTTGATGCAATTCGAA	1920
Db	1984	CATCGCGTTGAACTAGAGTACAGTACAGTAAACCTAGCCCACTATTTGATGCAATTCGAA <td>2043</td>	2043
QY	1921	ATTTGATGATGTTCTAGCAGTGAAGAGATTAAGAAAAAAGATCAACCAAA	1972
Db	2044	ATTTGATGATGTTCTAGCAGTGAAGAGATTAAGAAAAAAGATCAACCAAA	2095
RESULT 2			
LOCUS	BC030969	2090 bp mRNA linear HTC 13-JUN-2002	
DEFINITION	Homo sapiens, Similar to breast cancer 1, early onset, clone		
ACCESSION	IMAGE:4804551, mRNA.		
VERSION	BC030969		
KEYWORDS	BC030969.1 GI:21411299		
SOURCE	HTC.		
ORGANISM	human.		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 2090)		
JOURNAL	Strausberg, R. Direct Submission Submitted (03-JUN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>		
COMMENT	Contact: MGC help desk Email: <a href="mailto:cgabs-remail.nih.gov">cgabs-remail.nih.gov</a> Tissue Procurement: DCD/DTP cDNA Library Preparation: CIONTECH Laboratories, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (JLNL) Genome Center, Stanford University School of Medicine, Stanford, CA 94305 Web site: <a href="http://www.shgc.stanford.edu">http://www.shgc.stanford.edu</a> Contact: (Dickson, Mark) <a href="mailto:mdcpaxil.stanford.edu">mdcpaxil.stanford.edu</a> Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.		
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/JLNL at: <a href="http://image.jnl.nih.gov">http://image.jnl.nih.gov</a> Series: IRAL Plate: 41 Row: P Column: 11 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6552298 This clone has the following problem: frame shifted.		
SOURCE	Location/Qualifiers 1..2090 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:4804551" /tissue_type="Prostate, adenocarcinoma." /clone_id="NIH_MGC_60" /lab_host="DH10B" /note="Vector: pDNR-LIB"		
BASE COUNT	754 a 385 c 444 g 507 t		
ORIGIN			
Query Match	35.0%; Score 1955; DB 11; Length 2090;		
Best Local Similarity	99.5%; Pred. No. 0;		
Matches 1961; Conservative	0; Mismatches 10; Indels 0; Gaps 0;		
QY	1	ATGGATTATTCCTGCTTCGCGTTGAGAGAGTACAAAATGTCATTATTCGATACACAAA	60
Db	120	ATGGATTATTCCTGCTTCGCGTTGAGAGAGTACAAAATGTCATTATTCGATACACAAA	179
QY	61	ATCTTAAGAGTCCCATCTGCTGGAATGTGATCAAGAACCTGCTCCACAAAGTGAGAC	120
Db	180	ATCTTAAGAGTCCCATCTGCTGGAATGTGATCAAGAACCTGCTCCACAAAGTGAGAC	239
QY	121	CACATATTTTGCAAAATTTTGCAATCTGCAAACTTCTCAACGAGAAGAAAGGCGCTTCACAG	180

Db	240	CACATATTTTCCAAATTTTGCATGCTGAACCTTGCACCCAGAAAGGGCCTTACACAG	299
Qy	181	TGTCCTTTATGTAAAGATGATATTAACCAAAAGAGCCTACAGAAGATGACAGATTTAGT	240
Db	300	TGTCCTTTATGTAAAGATGATATTAACCAAAAGAGCCTACAGAAGATGACAGATTTAGT	359
Qy	241	CAACTGTGTGAAGAGCTATTGAAAATATATTGTACTTTTCAGCTTGTGACACAGATTTGGAG	300
Db	360	CAACTGTGTGAAGAGCTATTGAAAATATATTGTACTTTTCAGCTTGTGACACAGATTTGGAG	419
Qy	301	TATGCAAAACAGCTATTAATTTTGGCAAAAAGGAAATTAACCTCCTGAAACATCTAAAAGAT	479
Db	420	TATGCAAAACAGCTATTAATTTTGGCAAAAAGGAAATTAACCTCCTGAAACATCTAAAAGAT	479
Qy	361	GAACTTTCATCATCCAAAGATATGGGCTACGAAACCGTGGCAAAAGACCTTCACAGAGT	420
Db	480	GAACTTTCATCATCCAAAGATATGGGCTACGAAACCGTGGCAAAAGACCTTCACAGAGT	539
Qy	421	GAACCCGAAAAATCCCTCCCTGACAGAAAACCAAGCTCAGTGCACAACTCTCAACCTTGA	480
Db	540	GAACCCGAAAAATCCCTCCCTGACAGAAAACCAAGCTCAGTGCACAACTCTCAACCTTGA	599
Qy	481	ACTGTGAGAACTCTGAGAGCAAAACAGCAGGATACAACTCAAAAAGAGCTGTCTACTT	540
Db	600	ACTGTGAGAACTCTGAGAGCAAAACAGCAGGATACAACTCAAAAAGAGCTGTCTACTT	659
Qy	541	GAATTTGGAGATCTGATTCCTTGAAGATACCTTATATAAGCAACCTTATTTGAGTGGGA	600
Db	660	GAATTTGGAGATCTGATTCCTTGAAGATACCTTATATAAGCAACCTTATTTGAGTGGGA	719
Qy	601	GATCAAGAATGTTTACAAATCACCCCTCAAGAAACCAAGGATGAATCAGTTTGGATTC	660
Db	720	GATCAAGAATGTTTACAAATCACCCCTCAAGAAACCAAGGATGAATCAGTTTGGATTC	779
Qy	661	GCAAAAAAGCGCTGTGGAATTTTCTGAGACGAGATGTAAACAATPACGAACATCATCA	720
Db	780	GCAAAAAAGCGCTGTGGAATTTTCTGAGACGAGATGTAAACAATPACGAACATCATCA	839
Qy	721	CCCACTAATAATGATTTGAAACCAACCTGAGAAGCGTGCAGCTGAGAGGCATCCAGAAA	780
Db	840	CCCACTAATAATGATTTGAAACCAACCTGAGAAGCGTGCAGCTGAGAGGCATCCAGAAA	899
Qy	781	TATCAGGATACCTCTGTTTCAAACTTGAGTGTGAGCCATGTGGCACAAATACATCATGCC	840
Db	900	TATCAGGATACCTCTGTTTCAAACTTGAGTGTGAGCCATGTGGCACAAATACATCATGCC	959
Qy	841	AGCTCATTTACAGCATGAGAACAGCAGTTTATTACTACTAAAGACAGAAATAAATGTAGAA	900
Db	960	AGCTCATTTACAGCATGAGAACAGCAGTTTATTACTACTAAAGACAGAAATAAATGTAGAA	1019
Qy	901	AAGCGTGAATTCCTTAATAAAGCAAAACAGCTTGCTTACGACAGGACCAACATPACAGA	960
Db	1020	AAGCGTGAATTCCTTAATAAAGCAAAACAGCTTGCTTACGACAGGACCAACATPACAGA	1079
Qy	961	TGGCGTGAAGTAAAGGAAACATGTAAATGATATGGGGGACTCCACACAGCAAAAAAGGTA	1020
Db	1080	TGGCGTGAAGTAAAGGAAACATGTAAATGATATGGGGGACTCCACACAGCAAAAAAGGTA	1139
Qy	1021	GATCTGAATGCTGATCCCGCTGTGTGAGAGAAAGAAATGGAATTAAGCAAAAACTGCATGC	1080
Db	1140	GATCTGAATGCTGATCCCGCTGTGTGAGAGAAAGAAATGGAATTAAGCAAAAACTGCATGC	1199
Qy	1081	TCAGAGAAATCCTAAGAGATCTAGAAAGATGTTCTCTTGATACACTAATAATAGCAGATTACG	1140
Db	1200	TCAGAGAAATCCTAAGAGATCTAGAAAGATGTTCTCTTGATACACTAATAATAGCAGATTACG	1259
Qy	1141	AAATGTAATGAGTGGTTTTCACAGAAGGATGAACCTGTAGGTTGTGATGACTCACATGAT	1200
Db	1260	AAATGTAATGAGTGGTTTTCACAGAAGGATGAACCTGTAGGTTGTGATGACTCACATGAT	1319
Qy	1201	GGGGAGTCTGAATCAAAATGCCAAAGTAGCTGATGTATTGGACGTTCTPAAATGAGTAGAT	1260

Db 1330 GGGAGCTCTGAATCAAAATCCAAAGTAGCTGATGTATTGACGCTCTAAATGAGTAGAT 1379  
OY 1261 GAATATTTCTGTTCTCAGAGAAAATAGACTTACTGCGCAGTATCCTCATGAGGCTTTA 1320  
Db 1380 GAATATTTCTGTTCTCAGAGAAAATAGACTTACTGCGCAGTATCCTCATGAGGCTTTA 1439  
OY 1321 ATATGTAAAAGTAGAAAGAGTTCACCTCAATCAGTAGAGAGTAAATTTGAAGCAAAATA 1380  
Db 1440 ATATGTAAAAGTAGAAAGAGTTCACCTCAATCAGTAGAGAGTAAATTTGAAGCAAAATA 1499  
OY 1381 TTGGGAAAACCTATCGAAGAGCAACCTCCCACTTAAGCCATGTAAGTGAAGT 1440  
Db 1500 TTGGGAAAACCTATCGAAGAGCAACCTCCCACTTAAGCCATGTAAGTGAAGT 1559  
OY 1441 CTAAATATAGAGACATTTTCTACTGAGCAGACAGATATATCAAGAGGCTCCCTCACAAT 1500  
Db 1560 CTAAATATAGAGACATTTTCTACTGAGCAGACAGATATATCAAGAGGCTCCCTCACAAT 1619  
OY 1501 AAATTTAAAGCGTAAAGAGAGACCTACATCAGGCTTCATCTGAGAGATTTTATCAAGATA 1560  
Db 1620 AAATTTAAAGCGTAAAGAGAGACCTACATCAGGCTTCATCTGAGAGATTTTATCAAGATA 1679  
OY 1561 GGAGATTTGGAGCTTCAAAAGAGCTCCTGAATATGATTAATCAGAGAGTAAACCAAGGAG 1620  
Db 1680 GGAGATTTGGAGCTTCAAAAGAGCTCCTGAATATGATTAATCAGAGAGTAAACCAAGGAG 1739  
OY 1621 CAGAAATGCTCAGATGATGATATTTACTAATAGTGTCTCATGAGATTAACCAAGAGTGTAT 1680  
Db 1740 CAGAAATGCTCAGATGATGATATTTACTAATAGTGTCTCATGAGATTAACCAAGAGTGTAT 1799  
OY 1681 TCTATTTAGATGAGAAAAATCTCTAACCATATGATCACTGCAAAAAGAGATCTGCTTC 1740  
Db 1800 TCTATTTAGATGAGAAAAATCTCTAACCATATGATCACTGCAAAAAGAGATCTGCTTC 1859  
OY 1741 AAAAGCAAGCTGAAACCTTAAGCAGCAGTATTAAGCAATATGAGTAACTCGAATTAATATC 1800  
Db 1860 AAAAGCAAGCTGAAACCTTAAGCAGCAGTATTAAGCAATATGAGTAACTCGAATTAATATC 1919  
OY 1801 CACAATTTCAAAAGACCTTAAGAAATAGAGTGTGAGAGAGAGTCTTCTACAGGCAATAT 1860  
Db 1920 CACAATTTCAAAAGACCTTAAGAAATAGAGTGTGAGAGAGAGTCTTCTACAGGCAATAT 1979  
OY 1861 CATGGCTTTGAATCTGTAGTCTAGTAAGTAATCTAAGCCCTTAATTTGTACTGATTTGCA 1920  
Db 1980 CATGGCTTTGAATCTGTAGTCTAGTAAGTAATCTAAGCCCTTAATTTGTACTGATTTGCA 2039  
OY 1921 ATTGATAGTGTCTTACAGAGTGAAGATTAAGAAAAAGTAAACCAACCA 1971  
Db 2040 ATTGATAGTGTCTTACAGAGTGAAGATTAAGAAAAAGTAAACCAACCA 2090

RESULT 3  
LOCUS B0068830 962 bp mRNA linear EST 02-APR-2002  
DEFINITION AGENCOURT\_6740238 NIH\_MGC\_47 Homo sapiens cDNA clone IMAGE:5802685  
5', mRNA sequence.  
ACCESSION B0068830  
VERSION B0068830.1 GI:19897888  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 962)  
AUTHORS NIH-MGC <http://mgi.nhl.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LCM2037 row: h column: 14  
High quality sequence stop: 700.  
Location/Qualifiers

## FEATURES

source

1. 962

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5802685"  
/clone\_lib="NIH\_MGC\_47"  
/tissue\_type="neuroblastoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: brain; Vector: pOTB7; Site\_1: XhoI; Site\_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACAGG(G). Size selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."

BASE COUNT 302 a 195 c 235 g 229 t 1 others  
ORIGIN

Query Match 15.4%; Score 861.4; DB 14; Length 962;  
Best Local Similarity 96.2%; Pred. No. 1.7e-191;

Matches 890; Conservative 0; Mismatches 33; Indels 2; Gaps 1;

OY 4332 TCAGAACAAACACATCGAAAAAGCAGTATTAACCTCACAGAAAGTAGTATACC 4391  
Db 15 TTGGCAGCAGGACACATCGAAAAAGCAGTATTAACCTCACAGAAAGTAGTATACC 74  
OY 4392 TATTAACCCAGATCCGAAGAGCCTTCTGCTGACAGTTTGAGTGTCTGAGATAGTTC 4451  
Db 75 TATTAACCCAGATCCGAAGAGCCTTCTGCTGACAGTTTGAGTGTCTGAGATAGTTC 134  
OY 4452 TACCACTAAAAATATAAGAACAGAGAGTGAAGTGCATCCCTCTTAATGCCATCAT 4511  
Db 135 TACCACTAAAAATATAAGAACAGAGAGTGAAGTGCATCCCTCTTAATGCCATCAT 194  
OY 4512 AGATGATAGCTGTATGATGACACAGTTGCTGTGGAGTCTTACAGATTAAGTCCATC 4571  
Db 195 AGATGATAGCTGTATGATGACACAGTTGCTGTGGAGTCTTACAGATTAAGTCCATC 254  
OY 4572 TCAGAGAGAGCTCATTAAGGTTGTTGATGTGAGAGAGCAAGCTGGAAGAGTGGGCC 4631  
Db 255 TCAGAGAGAGCTCATTAAGGTTGTTGATGTGAGAGAGCAAGCTGGAAGAGTGGGCC 314  
OY 4632 ACAGATTTGACGGAACATCTTACTTGCACAGGCAAGATCTAGAGGAGACCCCTTACT 4691  
Db 315 ACAGATTTGACGGAACATCTTACTTGCACAGGCAAGATCTAGAGGAGACCCCTTACT 374  
OY 4692 GGAATCTGGAATCAGCCTTCTCTGTGATGACCTTAATCTGATCTTCTGAAGCAGAGC 4751  
Db 375 GGAATCTGGAATCAGCCTTCTCTGTGATGACCTTAATCTGATCTTCTGAAGCAGAGC 434  
OY 4752 CCCAGAGTACGCTCGTGTGGCAACATACCATCTTCAACCTGTGATGAAGTTCCCA 4811  
Db 435 CCCAGAGTACGCTCGTGTGGCAACATACCATCTTCAACCTGTGATGAAGTTCCCA 494  
OY 4812 ATTGAAGTTGACAGATCTGCCAGAGTCCAGAGTCCCTCATACTACTGATCTGCTGG 4871  
Db 495 ATTGAAGTTGACAGATCTGCCAGAGTCCAGAGTCCCTCATACTACTGATCTGCTGG 554  
OY 4872 GTATATATGCAATGGAAGAGTGTGAGCAGGAGAGACCAATTTGACAGCTTCAACAGA 4931  
Db 555 TTATATATCAAAAGAGAGAGTGTGAGCAGGAGAGCAAAATTTGACAGCTTCAACAGA 614  
OY 4932 AAGGCTCAACAAAGAAATGTCATGAGTGTGCTGAGCTGACCCAGAGAAATTTATGCT 4991  
Db 615 AAGGCTCAACAAAGAAATGTCATGAGTGTGCTGAGCTGACCCAGAGAAATTTATGCT 674

QY 4992 CGGTGACAGTTGGCCAGAAAACACACATCACTTAACTAATTAATTAAGAGAC 5051  
| | | | |  
Db 675 CGGTGACAGTTGGCCAGAAAACACACATCACTTAACTAATTAATTAAGAGAC 734  
| | | | |  
QY 5052 TACTCATGTTGTTATGAAGACAGTCTGAGTTGTGTGAGACGACACTGAATTTT 5111  
| | | | |  
Db 735 TACTCATGTTGTTATGAAGACAGTCTGAGTTGTGTGAGACGACACTGAATTTT 794  
| | | | |  
QY 5112 TCTAGAAATTCGGGAGGAAATGGTAGTATTTCTGAGTGAAGTGAATTTAA 5171  
| | | | |  
Db 795 TCTAGAAATTCGGGAGGAAATGGTAGTATTTCTGAGTGAAGTGAATTTAA 854  
| | | | |  
QY 5172 AGAAAGAAAAATGCTGAATGAGCATGATTTTGAAGTCAGAGAGATGTGG--TCAATGGA 5229  
| | | | |  
Db 855 AGAAAGAAAAATGCTGAATGAGCATGATTTTGAAGTCAGAGAGATGTGGTCAATGGA 914  
| | | | |  
QY 5230 AGAAAGAAAAATGCTGAATGAGCATGATTTTGAAGTCAGAGAGATGTGGTCAATGGA 914  
| | | | |  
Db 915 AGAAAGAAAAATGCTGAATGAGCATGATTTTGAAGTCAGAGAGATGTGGTCAATGGA 914  
| | | | |

RESULT 4  
BM452288  
LOCUS 1089 bp mRNA linear EST 05-FEB-2002  
DEFINITION AGENCOURT\_6386302 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:5526666  
5', mRNA sequence.  
ACCESSION BM452288  
VERSION BM452288.1 GI:18501328  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 1089)  
AUTHORS NIH-MGC http://mgi.ncl.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-r@mail.nih.gov  
Tissue Procurement: ATCC/DCMP/DMP  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.lnl.gov  
Plate: LLM12200 row: k column: 19  
High quality sequence stop: 667.  
Location/Qualifiers  
1. 1089

FEATURES  
Source  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5526666"  
/clone\_id="NIH\_MGC\_72"  
/tissue\_type="melanotic melanoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: skin; Vector: PCMV-SPORE6; Site: 1; Not:  
Site: 2; Sali: Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 2 kb. Library constructed by Life  
Technologies."  
BASE COUNT 389 a 212 c 248 g 240 t  
ORIGIN

Query Match 15.1%; Score 843; DB 13; Length 1089;  
Best Local Similarity 96.1%; Pred. No. 3.8e-187;  
Matches 929; Conservative 0; Mismatches 30; Indels 8; Gaps 6;

QY 1806 TTCAAAAGCACTTAAGAAATGAGTGTGAGAGAGAGTCTTCTACAGCATATTCATG 1865  
| | | | |  
Db 13 TTCAAAAGCACTTAAGAAATGAGTGTGAGAGAGAGTCTTCTACAGCATATTCATG 72  
| | | | |  
QY 1866 GCTTGAACGTGATGCTGATGAATCTAGCCCACTAATTTAGTGAATTTGA 1925  
| | | | |

Db 73 GCTTGAACGTGATGCTGATGAATCTAGCCCACTAATTTAGTGAATTTGA 132  
| | | | |  
QY 1926 TAGTGTCTTACAGTGAAGATGAAGAAAGAAAGTACAAACCAATGCCATGAGCA 1985  
| | | | |  
Db 133 TAGTGTCTTACAGTGAAGATGAAGAAAGTACAAACCAATGCCATGAGCA 192  
| | | | |  
QY 1986 CAGCAGAAACCTACACTGATGAAGTAAAGAACTGCAACTGGAGCAGAGAGATTA 2045  
| | | | |  
Db 193 CAGCAGAAACCTACACTGATGAAGTAAAGAACTGCAACTGGAGCAGAGAGATTA 252  
| | | | |  
QY 2046 CAAGCCAAATGACAGACAAAGTAAAGACATGACATGATCTTCCAGAGCTGAAGTT 2105  
| | | | |  
Db 253 CAAGCCAAATGACAGACAAAGTAAAGACATGACATGATCTTCCAGAGCTGAAGTT 312  
| | | | |  
QY 2106 AACAAATGACACCTGTTCTTTTACTAGTGTTCAAATACAGTGAATTAAGATTTGT 2165  
| | | | |  
Db 313 AACAAATGACACCTGTTCTTTTACTAGTGTTCAAATACAGTGAATTAAGATTTGT 372  
| | | | |  
QY 2166 CAATCTAGACCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2225  
| | | | |  
Db 373 CAATCTAGACCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 432  
| | | | |  
QY 2226 TAATGCTGAAGACCCCAAGATCTCATGTTTACGTGAGAGAGAGAGAGAGAGAG 2285  
| | | | |  
Db 493 ATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 552  
| | | | |  
QY 2286 ATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2345  
| | | | |  
Db 493 ATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 552  
| | | | |  
QY 2346 TATCTGCTTACTGAGAGTATGACACTTACAGTGAAGAGAGAGAGAGAGAGAGAG 2405  
| | | | |  
Db 553 TATCTGCTTACTGAGAGTATGACACTTACAGTGAAGAGAGAGAGAGAGAGAGAG 612  
| | | | |  
QY 2406 GAGTCACTGTGACAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2465  
| | | | |  
Db 613 GAGTCACTGTGACAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 672  
| | | | |  
QY 2466 TAGAAATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2525  
| | | | |  
Db 673 TAGAAATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 732  
| | | | |  
QY 2526 AACAGAGATGAAGT--GGAAGAAAGTGAAGTGTGCTGATTTCCAGAAATCATTTCA 2584  
| | | | |  
Db 733 AACAGAGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 792  
| | | | |  
QY 2585 AGGTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2644  
| | | | |  
Db 793 A-GTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 850  
| | | | |  
QY 2645 GTGCAACATTTCTGCGCCACTGCTGGTGC--TTAAGAAACAAAGTCCAAAGTCACTTTT 2703  
| | | | |  
Db 851 GTGCAACATTTCTGCGCCACTGCTGGTGCCTTTAAGAAACAAAGTCCAAAGTCCCTTTT 910  
| | | | |  
QY 2704 GAATGTGACAAAAA---GGAAGAAATCAAGAGAGAGATGATCA--TATCAAGCTGT 2759  
| | | | |  
Db 911 GAATGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 970  
| | | | |  
QY 2760 ACAGACA 2766  
| | | | |  
Db 971 AACGAAA 977  
| | | | |

RESULT 5  
AUI22476  
LOCUS 845 bp mRNA linear EST 01-AUG-2002  
DEFINITION AUI22476 MAMMA1 Homo sapiens cDNA clone MAMMA1002447 5', mRNA  
sequence.  
ACCESSION AUI22476  
VERSION AUI22476.1 GI:10937746  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Unpublished (2000)  
Contact: Takao Isogai  
Genomics Laboratory  
Helix Research Institute  
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
Tel: 81-438-52-3975  
Fax: 81-438-52-3986  
Email: genomics@hri.co.jp  
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix  
Research Institute; cDNA library construction: Department of  
Virology, Institute of Medical Science, University of Tokyo, and  
Helix Research Institute.

FEATURES  
Source  
1. 845  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="MAMMA1002447"  
/clone\_lib="MAMMA1"  
/cissue\_type="mammary gland"  
/note="Vector: PME18SFL3"

BASE COUNT 279 a 172 c 191 g 200 t 3 others  
ORIGIN

Query Match 14.5%; Score 811.8; DB 9; Length 845;  
Best Local Similarity 99.1%; Pred. No. 7.5e-180;  
Matches 836; Conservative 0; Mismatches 5; Indels 3; Gaps 2;

QY 3700 GTAACAAATATACCTTCTAGTCTAGTACGATAGCAACCGCTTCTACCGAGTGTCTCT 3759  
DB 1 GTAACAAATATACCTTCTAGTCTAGTACGATAGCAACCGCTTCTACCGAGTGTCTCT 60  
QY 3760 AAGAACAAGAGAGAGATTTATATCATTTGAAGATAGCTTAATGACTGCTCAATACCG 3819  
DB 61 AAGAACAAGAGAGAGATTTATATCATTTGAAGATAGCTTAATGACTGCTCAATACCG 120  
QY 3820 GTAATATTTGGCAAAAGGATCTCAGAAATCATCACCTTATAGTGAAGAAACAAATGTTCTGT 3879  
DB 121 GTAATATTTGGCAAAAGGATCTCAGAAATCATCACCTTATAGTGAAGAAACAAATGTTCTGT 180  
QY 3880 AACTTGTCTTCTTACAGAGCACTGATGTAATGGAAGACTTGTCACTCAATACAAACACCCAG 3939  
DB 181 AACTTGTCTTCTTACAGAGCACTGATGTAATGGAAGACTTGTCACTCAATACAAACACCCAG 240  
QY 3940 GATCCTTCTTCTTATGTTCTTCCAAACAATGAGGATGCTGAAAGCCAGGGAGTT 3999  
DB 241 GATCCTTCTTCTTATGTTCTTCCAAACAATGAGGATGCTGAAAGCCAGGGAGTT 300  
QY 4000 GGTCTGAGTACAGAAATTTGTTTCAGATGATGAAGAAAGAGAAACGGCTTGAAGAA 4059  
DB 301 GGTCTGAGTACAGAAATTTGTTTCAGATGATGAAGAAAGAGAAACGGCTTGAAGAA 360  
QY 4060 AATATCAAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4119  
DB 361 AATATCAAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420  
QY 4120 ACTGAACAAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4179  
DB 421 ACTGAACAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480  
QY 4180 ACTGACAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4239  
DB 481 ACTGACAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540  
QY 4240 CTAGAAGCTGTGTAGACAGATGGAGCCAGCTTCTTAACAGTACCTCTTCATCATATA 4299  
DB 541 CTAGAAGCTGTGTAGACAGATGGAGCCAGCTTCTTAACAGTACCTCTTCATCATATA 600

QY 4300 AGTGAATCTTCTGCTCCCTTGGAGACCTGGCAAAATCCGAACAAAGACATCAGAAAAAGCA 4359  
DB 601 AGTGAATCTTCTGCTCCCTTGGAGACCTGGCAAAATCCGAACAAAGACATCAGAAAAAGCA 660  
QY 4360 GTATTAATCTTACAGAGAAAGTAGTGAATACCCCTATTAAGCCAGAAATCCAGAAAGCTTCT 4419  
DB 661 GTATTAATCTTACAGAGAAAGTAGTGAATACCCCTATTAAGCCAGAAATCCAGAAAGCTTCT 720  
QY 4420 GCTGACAGAGTTGAGTGTCTGAGATAGTCTTACCAAGTAAATTAAGAACCCAGCA--G 4477  
DB 721 GCTGACAGAGTTGAGTGTCTGAGATAGTCTTACCAAGTAAATTAAGAACCCAGCAAGT 780  
QY 4478 TGGAAAGGTCA-TCCCTCTTAATAGCCATCATTAATGATGATAGGATGATGATGATGATGAT 4536  
DB 781 TGGAAAGGTCAATCCCTCTTCTTAATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
QY 4537 TGCT 4540  
DB 841 TGCT 844

RESULT 6  
AUI42729 783 bp mRNA linear EST 05-AUG-2002  
LOCUS AUI42729 y79AA1 Homo sapiens cDNA clone y79AA1000792 5', mRNA  
DEFINITION sequence.  
ACCESSION AUI42729 GI:11004250  
VERSION AUI42729.1  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Unpublished (2000)  
Contact: Takao Isogai  
Genomics Laboratory  
Helix Research Institute  
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
Tel: 81-438-52-3975  
Fax: 81-438-52-3986  
Email: genomics@hri.co.jp  
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix  
Research Institute; cDNA library construction: Department of  
Virology, Institute of Medical Science, University of Tokyo, and  
Helix Research Institute.

FEATURES  
Source  
1. 783  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="y79AA1000792"  
/clone\_lib="y79AA1"  
/cell\_type="retinoblastoma"  
/note="Vector: PME18SFL3"

BASE COUNT 255 a 146 c 185 g 194 t 3 others  
ORIGIN

Query Match 13.6%; Score 761.8; DB 9; Length 783;  
Best Local Similarity 99.2%; Pred. No. 4.2e-168;  
Matches 774; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 3460 TTAGATGATGGTGAATAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3519  
DB 1 TTAGATGATGGTGAATAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60  
QY 3520 TCTGCTTTTTCAGCAAAAGCTCCAGAAAGAGAGCTTAGCAGAGAGCTTAGCCCTTTC 3579  
DB TCTGCTTTTTCAGCAAAAGCTCCAGAAAGAGAGCTTAGCAGAGAGCTTAGCCCTTTC 3579



Db 61 TGTGCTGTTTTAGCAAAAAGCGTCCAGAAAGAGAGCTTAGCAGAGCTTACGCTTTC 120  
 QY 3580 ACCCATACACATTTGGCTCAGGTTACCGAAGAGGGGCCAAGAAATTAGAGTCTCAGAA 3639  
 Db 121 ACCCATACACATTTGGCTCAGGTTACCGAAGAGGGGCCAAGAAATTAGAGTCTCAGAA 180  
 QY 3640 GAGAACTATCTAGTAGAGATGAAGAGCTTCCCTGCTTCCAAACACTTGTATTGGTAA 3699  
 Db 181 GAGAACTATCTAGTAGAGATGAAGAGCTTCCCTGCTTCCAAACACTTGTATTGGTAA 240  
 QY 3700 GTAAACAATATACCTTCTCAGTCTCTAGCTAGCAGTACACCGTGTACCGAGTGTCTGCT 3759  
 Db 241 GTAAACAATATACCTTCTCAGTCTCTAGCTAGCAGTACACCGTGTACCGAGTGTCTGCT 300  
 QY 3760 AAGAACACAGAGAGAGATTTATTCATTCATTAAGAAATAGCTTAATGATGCTAGTAACCG 3819  
 Db 301 AAGAACACAGAGAGAGATTTATTCATTCATTAAGAAATAGCTTAATGATGCTAGTAACCG 360  
 QY 3820 GTAAATTTGGCAAGGATCTCAGAGAACATCCTTAGTGAGAAACAAATGTTCTGCT 3879  
 Db 361 GTAAATTTGGCAAGGATCTCAGAGAACATCCTTAGTGAGAAACAAATGTTCTGCT 420  
 QY 3880 AGCTGTTTCTTCTCAGTCTCAGTCTCAGTCTCAGTCTCAGTCTCAGTCTCAGTCTCAGTCT 3939  
 Db 421 AGCTGTTTCTTCTCAGTCTCAGTCTCAGTCTCAGTCTCAGTCTCAGTCTCAGTCTCAGTCT 480  
 QY 3940 GATCCTTTCTTCTCAGTCTCAGTCTCAGTCTCAGTCTCAGTCTCAGTCTCAGTCTCAGTCT 3999  
 Db 481 GATCCTTTCTTCTCAGTCTCAGTCTCAGTCTCAGTCTCAGTCTCAGTCTCAGTCTCAGTCT 540  
 QY 4000 GGTCTGAGTACAGAGATTTGTTTCAATGATGATGAAGAAAGAGAGGGGCTTGGAGAA 4059  
 Db 541 GGTCTGAGTACAGAGATTTGTTTCAATGATGATGAAGAAAGAGAGGGGCTTGGAGAA 600  
 QY 4060 AATATCAAGAGAGAGAGATTTGTTTCAATGATGATGAAGAGAGAGAGAGAGAGAGAGAGAG 4119  
 Db 601 AATATCAAGAGAGAGAGATTTGTTTCAATGATGATGAAGAGAGAGAGAGAGAGAGAGAGAG 660  
 QY 4120 AGTGAACAGAGAGAGAGATTTGTTTCAATGATGATGAAGAGAGAGAGAGAGAGAGAGAGAG 4178  
 Db 661 AGTGAACAGAGAGAGAGATTTGTTTCAATGATGATGAAGAGAGAGAGAGAGAGAGAGAGAG 720  
 QY 4179 CACTCAGCAGAGAGAGATTTGTTTCAATGATGATGAAGAGAGAGAGAGAGAGAGAGAGAG 4238  
 Db 721 CACTCAGCAGAGAGAGATTTGTTTCAATGATGATGAAGAGAGAGAGAGAGAGAGAGAGAG 780  
 RESULT 7  
 Bg681276 743 bp mRNA linear EST 01-MAY-2001  
 LOCUS 602627125F1 NCI\_CGAP\_Skn4 Homo sapiens cDNA clone IMAGE:4751887 5',  
 DEFINITION mRNA sequence.  
 ACCESSION Bg681276  
 VERSION Bg681276.1 GI:13912673  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 743)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-rt@mail.nih.gov  
 Tissue Procurement: James Cleaver, M.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLM10609 row: a column: 08

FEATURES High quality sequence stop: 741.  
 source Location/Qualifiers  
 1..743  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4751887"  
 /clone\_lib="NCI\_CGAP\_Skn4"  
 /tissue\_type="squamous cell carcinoma"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /note="Organ: skin; Vector: pCMV-Sport6; Site: 1: NotI;  
 Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
 Average insert size 1.5kb. Library constructed by Life  
 Technologies. Note: this is a NCI\_CGAP Library."  
 BASE COUNT 233 a 156 c 182 g 172 t  
 ORIGIN  
 Query Match 13.2%; Score 740.4; DB 12; Length 743;  
 Best Local Similarity 99.9%; Pred. No. 4,4e-163;  
 Matches 741; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Db 4605 GGAGCAACAGCTGGAAGAGTGGGCGACAGATTTGACGAAACATCTTACTTGGCAAG 4664  
 1 GGAGCAACAGCTGGAAGAGTGGGCGACAGATTTGACGAAACATCTTACTTGGCAAG 60  
 QY 4665 GCAAGATCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4724  
 Db 61 GCAAGATCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120  
 QY 4725 TGAATCTGATCTTCTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4784  
 Db 121 TGAATCTGATCTTCTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180  
 QY 4785 TTCAACCTCTGATGGAAGAGTGGGCGACAGATTTGACGAAACATCTTACTTGGCAAG 4844  
 Db 181 TTCAACCTCTGATGGAAGTGGGCGACAGATTTGACGAAACATCTTACTTGGCAAG 240  
 QY 4845 TGCTGCTCATCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4904  
 Db 241 TGCTGCTCATCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
 QY 4905 GAGGCAAGATTTGACAGCTTCAACAGAAAGAGTCAACAAAGAGTCCATGCTGCTGCTGCTGCT 4964  
 Db 301 GAGGCAAGATTTGACAGCTTCAACAGAAAGAGTCAACAAAGAGTCCATGCTGCTGCTGCTGCT 360  
 QY 4965 TGGCCGACCCCAAGAGATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5024  
 Db 361 TGGCCGACCCCAAGAGATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
 QY 5025 TTTAACTAATCTAATTTACTGAGAGAGATCTGATGATGATGATGATGATGATGATGATGATGAT 5084  
 Db 421 TTTAACTAATCTAATTTACTGAGAGAGATCTGATGATGATGATGATGATGATGATGATGATGAT 480  
 QY 5085 TGTGTGTGAGAGAGAGATTTATTTCTAGGAATTTGGGAGAGAAATGGGTAGTTAG 5144  
 Db 481 TGTGTGTGAGAGAGAGATTTATTTCTAGGAATTTGGGAGAGAAATGGGTAGTTAG 540  
 QY 5145 CTATTCTGGGAGACCAAGTCTAATTAAGAAAGAAATGCGATGATGATGATGATGATGATGATGAT 5204  
 Db 541 CTATTCTGGGAGACCAAGTCTAATTAAGAAAGAAATGCGATGATGATGATGATGATGATGATGAT 600  
 QY 5205 AGTCAGAGAGATGTGTCAATGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 5264  
 Db 601 AGTCAGAGAGATGTGTCAATGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 660  
 QY 5265 CCAGAGAGAGATTTCTCAGGGGCTGGAATTTCTGCTGATGAGGCGCTTACCACATAT 5324  
 Db 661 CCAGAGAGAGATTTCTCAGGGGCTGGAATTTCTGCTGATGAGGCGCTTACCACATAT 720  
 QY 5325 GCCCAGACATCACTGGAATGG 5346  
 Db 721 GCCCAGACATCACTGGAATGG 742

RESULT 8	
LOCUS	AU125312
DEFINITION	AU125312 NT2RM4 Homo sapiens cDNA clone NT2RM4001388 5' mRNA sequence.
ACCESSION	AU125312
VERSION	AU125312.1 GI:10950028
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S., Saito,K., Yamamoto,J., Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki,Y., Sugano,S. and Isogai,T.
TITLE	HRI human CDNA project (Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S., Saito,K., Yamamoto,J., Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki,Y., Sugano,S., Isogai,T.)
JOURNAL	Unpublished (2000)
COMMENT	Contact: Takao Isogai Genomics Laboratory Helix Research Institute 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan Tel.: 81-438-52-3975 Fax: 81-438-52-3986 Email: genomics@hri.co.jp HRI human CDNA project; 5'- & 3'-end one pass sequencing; Helix Research Institute; CDNA library construction; Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute. Location/Qualifiers 1..702 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="NT2RM4001388" /clone_1id="NT2RM4" /cell_type="teratocarcinoma" /note="Vector: pME18SFL3; mRNA from uninduced NT2 neuronal precursor cells"
FEATURES	
SOURCE	
BASE COUNT	235 a 142 c 163 g 159 t 3 others
ORIGIN	
Query Match	12.3%; Score 688; DB 9; Length 702;
Best Local Similarity	99.4%; Pred. No. 9,2e-151;
Matches 699; Conservative	0; Mismatches 3; Indels 1; Gaps 1;
Oy	3822 AATATTGGCAAAAGCATCTCAGAACATCACACCCTTAGTGTGAGGAACAATAATGTCGTCTAG 3881
Dd	1 AATATTGGCAAAAGCATCTCAGAACATCACACCCTTAGTGTGAGGAACAATAATGTCGTCTAG 60
Oy	3882 CTGTGTTTCTTCACAGTGTCAGTGAATTTGGAAGACTTGACTGCATTAACAAACCCACAGA 3941
Ddb	61 CTGTGTTTCTTCACAGTGTCAGTGAATTTGGAAGACTTGACTGCATTAACAAACCCACAGA 120
Oy	3942 TCCTTTCTTGAATGGTCTTCCCAACAATGAGCGCATGCTGTGAAGCCAGGAGTTGG 4001
Dd	121 TCCTTTCTTGAATGGTCTTCCCAACAATGAGCGCATGCTGTGAAGCCAGGAGTTGG 180
Oy	4002 TCTGAGTGACAAGGAATGTTTTCAGATGATGAAGAAAAGCAACGGCTTGAAGAAA 4061
Dd	181 TCTGAGTGACAAGGAATGTTTTCAGATGATGAAGAAAAGCAACGGCTTGAAGAAA 240
Oy	4062 TAATCAAGAGCAAAAGCATGATGATTCAAACTTAGGTGAAGCAGATTCGTGGGTGAGAG 4121
Dd	241 TAATCAAGAGCAAAAGCATGATGATTCAAACTTAGGTGAAGCAGATTCGTGGGTGAGAG 300
Oy	4122 TGAACAAGCGCTCTCTGAAGACTGCTCAGGCTATCTCTCAGATGATCATTTAACCAC 4181
Dd	301 TGAACAAGCGCTCTCTGAAGACTGCTCAGGCTATCTCTCAGATGATCATTTAACCAC 360
Oy	4182 TCAGCAGAGGATACCATGACAATACCTGATTAAGAGCTCCAGCAGGAATGGCTGAAC 4241

[illegible]

performed as previously described [Bonafide, Lennon & Soares (1996): Normalization and Subtraction: Two Approaches To Facilitate Gene Discovery. Genome Research 6: 791-806.

\*tag-SFQ=None found\*

BASE COUNT 173 a 141 c 128 g 295 t 2 others

Query Match 12.3%; Score 686.6; DB 12; Length 739;

Best Local Similarity 99.1%; Pred. No. 2e-150; Mismatches 6; Gaps 0;

Matches 689; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

1267 TCTGTTCTTCAGAGAAATAGACTTCTGCGCAGTATCTCATGAGCCTTAAATATGT 1326  
 1268 TCTGTTCTTCAGAGAAATAGACTTCTGCGCAGTATCTCATGAGCCTTAAATATGT 1326  
 1327 AAAAGTGAAGAGTCTTCTCAATATGATAGAGATTAATTTGAGACAAATATTGGG 1386  
 638 AAAAGTGAAGAGTCTTCTCAATATGATAGAGATTAATTTGAGACAAATATTGGG 579  
 1387 AAAAGTGAAGAGTCTTCTCAATATGATAGAGATTAATTTGAGACAAATATTGGG 1446  
 578 AAAAGTGAAGAGTCTTCTCAATATGATAGAGATTAATTTGAGACAAATATTGGG 519  
 1447 ATAGAGCATTTTGTACTGAGCCACAGATTAATACAGAGCCTTCCCTCAATAATATTA 1506  
 518 ATAGAGCATTTTGTACTGAGCCACAGATTAATACAGAGCCTTCCCTCAATAATATTA 459  
 1507 AAGCGTAAAGAGAGCTTCAATATGATAGAGATTAATTTGAGACAAATATTGGG 1566  
 458 AAGCGTAAAGAGAGCTTCAATATGATAGAGATTAATTTGAGACAAATATTGGG 399  
 1567 TTGGCGATTCAGAGAGAGCTTCAATATGATAGAGATTAATTTGAGACAAATATTGGG 1626  
 398 TTGGCGATTCAGAGAGAGCTTCAATATGATAGAGATTAATTTGAGACAAATATTGGG 339  
 1627 GGTCAGATGATATTAATTAATGATAGAGATTAATTTGAGACAAATATTGGG 1686  
 338 GGTCAGATGATATTAATTAATGATAGAGATTAATTTGAGACAAATATTGGG 279  
 1687 CAGATGAGAGAGAGCTTCAATATGATAGAGATTAATTTGAGACAAATATTGGG 1746  
 278 CAGATGAGAGAGAGCTTCAATATGATAGAGATTAATTTGAGACAAATATTGGG 219  
 1747 AAGCGTAAAGAGAGCTTCAATATGATAGAGATTAATTTGAGACAAATATTGGG 1806  
 218 AAGCGTAAAGAGAGCTTCAATATGATAGAGATTAATTTGAGACAAATATTGGG 159  
 1807 TCAAAAGAGAGAGCTTCAATATGATAGAGATTAATTTGAGACAAATATTGGG 1866  
 158 TCAAAAGAGAGAGCTTCAATATGATAGAGATTAATTTGAGACAAATATTGGG 99  
 1867 CTTGAAGTATGATAGAGAGCTTCAATATGATAGAGATTAATTTGAGACAAATATTGGG 1926  
 98 CTTGAAGTATGATAGAGAGCTTCAATATGATAGAGATTAATTTGAGACAAATATTGGG 39  
 1927 AGTTGTTCTAGAGAGAGAGATTAATTTGAGACAAATATTGGG 1961  
 38 AGTTGTTCTAGAGAGAGAGATTAATTTGAGACAAATATTGGG 4

RESULT 10  
 BM800251 987 bp mRNA linear EST 05-MAR-2002  
 LOCUS AGENCOURT\_6415998 NIH\_MGC\_71 Homo sapiens cDNA clone IMAGE:5531286  
 DEFINITION 5' mRNA sequence.  
 ACCESSION BM800251  
 VERSION BM800251.1 GI:19117074  
 KEYWORDS EST.  
 SOURCE human  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 987)  
 AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgabs-r@mail.nih.gov  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.lnl.gov>  
 Plate: LLM12212 row: 1 column: 07  
 High quality sequence stop: 692.

## FEATURES

## source

location/Qualifiers

1..987

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:5531286"

/clone\_lib="NIH\_MGC-71"

/issue\_type="telomysarcoma"

/lab\_host="DH10B (phage-resistant)"

/note="Organ: uterus; Vector: pCMV-SPORT6; Site: 1; NotI;  
 Site: 2; SalI; Cloned unidirectionally. Primer: Oligo dT.  
 Average insert size 2.1 kb."

BASE COUNT 342 a 181 c 222 g 238 t 4 others

## ORIGIN

Query Match 12.3%; Score 686.4; DB 14; Length 987;

Best Local Similarity 85.8%; Pred. No. 2.5e-150;

Matches 874; Conservative 0; Mismatches 14; Indels 131; Gaps 5;

279 TCACTTGACACAGCTTGGAGTATGCAACAGCTTAATTTGCAAAAAAGAAATTA 338  
 1 TAGCTTGACACAGCTTGGAGTATGCAACAGCTTAATTTGCAAAAAAGAAATTA 60  
 339 CTCTCTGCAACATCTTAAGATGAGTATCTTATCATCCAAAGATGAGGCTACGAAACCG 398  
 61 CTCTCTGCAACATCTTAAGATGAGTATCTTATCATCCAAAGATGAGGCTACGAAACCG 120  
 399 TGCAGAAAGAGCTTCAACAGAGTGAACCGAAATCTCTCTGACAGAAACAGTCTCAG 458  
 121 TGCAGAAAGAGCTTCAACAGAGTGAACCGAAATCTCTCTGACAGAAACAGTCTCAG 177  
 459 TGTCCACACTCTAACCTTGAAGTGTGAGACTGTGAGAGACAAACAGCGATACAC 518  
 178 TGTCCACACTCTAACCTTGAAGTGTGAGACTGTGAGAGACAAACAGCGATACAC 237  
 519 TCAAAAGAGAGCTTCAACAGAGTGAAGTGTGAGACTGTGAGAGACAAACAGCGATAC 578  
 238 TCAAAAGAGAGCTTCAACAGAGTGAAGTGTGAGACTGTGAGAGACAAACAGCGATAC 264  
 579 GGCACACTTATGACAGTGTGGAGATCAAGAAATGTTACAAATCACCCCTCAAGGAACAG 638  
 265 -----GGCTGCTTGTGAATTTCTGAGACGATGT 264  
 639 GCATGAATACAGTTGGATTTCGAAAAAGAGCTGTTGTGAATTTCTGAGACGATGT 698  
 265 -----GGCTGCTTGTGAATTTCTGAGACGATGT 294  
 699 AACCAATACAGATCATCAACCCAGTATATGATTTGAACACACAGAGAGGTC 758  
 295 AACCAATACAGATCATCAACCCAGTATATGATTTGAACACACAGAGAGGTC 354  
 759 AGCTGAGAGAGATCCAGAAAGATATCAGAGTGTGTTTCAAACTTCATGAGAGCC 818  
 355 AGCTGAGAGAGATCCAGAAAGATATCAGAGTGTGTTTCAAACTTCATGAGAGCC 414  
 819 ATGTGGCAAAATACTCATGCTCAGCTCATTTACAGCATGGAACAGCAGTTTACTAC 878  
 415 ATGTGGCAAAATACTCATGCTCAGCTCATTTACAGCATGGAACAGCAGTTTACTAC 474

QY 879 TAAAGCAGAAATGAAATGTAGAAAAGCGTGAATTTCTGTAAATTAAGCAACAGCTGGCTT 938  
 DB 475 TAAAGCAGAAATGAAATGTAGAAAAGCGTGAATTTCTGTAAATTAAGCAACAGCTGGCTT 534  
 QY 939 AGCAAGGAGCAACATACAGATGGGCTGGAAGTAGGAACATGTAATGATAGCGGAC 998  
 DB 535 AGCAAGGAGCAACATACAGATGGGCTGGAAGTAGGAACATGTAATGATAGCGGAC 594  
 QY 999 TCCGACGACAGAAAAAAGTATGATCTGATCCCTGTGTGAGAGAAAAGATG 1058  
 DB 595 TCCGACGACAGAAAAAAGTATGATCTGATCCCTGTGTGAGAGAAAAGATG 654  
 QY 1059 GAATTAAGCAGAAATGCGATGCTCAGAGAACTCTAGAGATCTGAAGATGTTCTTGAT 1118  
 DB 655 GAATTAAGCAGAAATGCGATGCTCAGAGAACTCTAGAGATCTGAAGATGTTCTTGAT 714  
 QY 1119 AACCTAATATGACATTCAGAAAGTTAATGAGGGTTTCCAGAGATGATGATGTT 1178  
 DB 715 AACCTAATATGACATTCAGAAAGTTAATGAGGGTTTCCAGAGATGATGATGTT 774  
 QY 1179 AGGTTCGATGACATTCAGATGGGAGCTGGAATCAATGCCAAAGTACGT -GATGAT 1237  
 DB 775 AGGTTCGATGACATTCAGATGGGAGCTGGAATCAATGCCAAAGTACGTGATGAT 834  
 QY 1238 TGGACGTT-CTTAATGAGGTAGATG---AATATTTGTTCTTCAGAGAAAATAGACTT 1292  
 DB 835 TGGACGTTCTTAATGAGGTAGATGATGATTTCTGTTCTTCAGAGAAAATAGACTT 893

RESULT 11  
 LOCUS B0422380 899 bp mRNA linear EST 23-MAY-2002  
 DEFINITION AGNCOURT\_7802085 NIH\_MGC\_92 Homo sapiens cDNA clone IMAGE:6042052  
 B0422380  
 VERSION B0422380.1 GI:21117695  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 899)  
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: csaabp@rmail.nih.gov

Tissue Procurement: ATCC  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 plate: LLNL3280 row: n column: 05  
 High quality sequence stop: 597.  
 Location/Qualifiers

## FEATURES

source

1. 899  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6042052"  
 /clone\_lib="NIH\_MGC\_92"  
 /tissue\_type="embryonal carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: testis; Vector: pCMV-SPORT6; Site: 1: NotI; Site: 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH-MGC Library."  
 BASE COUNT 291 a 182 c 192 g 230 t 4 others

Query Match

11.9%; Score 666.6; DB 14; Length 899;

Best Local Similarity 97.5%; Pred. No. 1,1e-145;  
 Matches 740; Conservative 0; Mismatches 11; Indels 8; Gaps 6;  
 QY 1 ATGATTTATCTGCTCTCCGTTGAAGAACTATCAAAATGCTAATTAATGATACAGAAA 60  
 DB 100 ATGATTTATCTGCTCTCCGTTGAAGAACTATCAAAATGCTAATTAATGATACAGAAA 159  
 QY 61 ATCTTAGAGTGTCCTCATCTGCTGGAGTTGATCAAGAACTCTCTCACAAAGTGTAC 120  
 DB 160 ATCTTAGAGTGTCCTCATCTGCTGGAGTTGATCAAGAACTCTCTCACAAAGTGTAC 219  
 QY 121 CACATATTTTGGCAATTTTTCATGCTGAACCTTCTCAACCAAGAAAGAGGCTTCACAG 180  
 DB 220 CACATATTTTGGCAATTTTTCATGCTGAACCTTCTCAACCAAGAAAGAGGCTTCACAG 279  
 QY 181 TGTCTTTATGTAATATGATATTAACCAAGAGAGCTCAGAAAGATGATTTAGT 240  
 DB 280 TGTCTTTATGTAATATGATATTAACCAAGAGAGCTCAGAAAGATGATTTAGT 339  
 QY 241 CAACCTGTTGAAGAGCTATGAAATATCATTTTGTCTTTCAGTTGACAGAGTTTGAG 300  
 DB 340 CAACCTGTTGAAGAGCTATGAAATATCATTTTGTCTTTCAGTTGACAGAGTTTGAG 399  
 QY 301 TATGCAACAGCTATATTTTTCAGAAAAGCAAAATTAATCTCTGTAACATCTTAAAGAT 360  
 DB 400 TATGCAACAGCTATATTTTTCAGAAAAGCAAAATTAATCTCTGTAACATCTTAAAGAT 459  
 QY 361 GAAGTTTCATCATCAAAAGATGAGGCTACAGAAACCGTGCAGAAAGATCTTACAGAGT 420  
 DB 460 GAAGTTTCATCATCAAAAGATGAGGCTACAGAAACCGTGCAGAAAGATCTTACAGAGT 519  
 QY 421 GAACCCGAAATCTCTTCGAGAAACAGCTCAGTGTCAACCTCTTAACCTTGA 480  
 DB 520 GAACCCGAAATCTCTTCGAGAAACAGCTCAGTGTCAACCTCTTAACCTTGA 576  
 QY 481 ACTGTGAGAACTCTAGAGCAAGACGAGGATTAACCTCAAAAGAGCTGTCTACATT 540  
 DB 577 ACTGTGAGAACTCTAGAGCAAGACGAGGATTAACCTCAAAAGAGCTGTCTACATT 636  
 QY 541 GAATGGGATGATCTTCTTGAAGATACCTTATTAAGCAATTTTCAGGTGAGG 600  
 DB 637 GAATGGGATGATCTTCTTGAAGATACCTTATTAAGCAATTTTCAGGTGAGG 696  
 QY 601 GATCAAGATTTGTAACAATCAACCCCTCAAGAACCA -GGATGAAATCAGTTGGATTG 659  
 DB 697 GATCAAGATTTGTAACAATCAACCCCTCAAGAACCA -GGATGAAATCAGTTGGATTG 756  
 QY 660 TGCAGAAAAGGCTGC-TTGTGAATTTTCTGAGACGAGATGTACAA -ATACTGAACATC-A 716  
 DB 757 TGCAGAAAAGGCTGCTTTGTGAATTTTCTGAGACGAGATGTACAA -ATACTGAACATC-A 816  
 QY 717 TCACCCAGTATATATGA-TTTGAACACCACTGGAAGC 754  
 DB 817 TCACCCAGTATATATGA-TTTGAACACCACTGGAAGC 855

## RESULT 12

BF791668

LOCUS BF791668

DEFINITION BF791668 900 bp mRNA linear EST 12-JAN-2001

VERSION BF791668

ACCESSION BF791668.1 GI:12096722

KEYWORDS mRNA sequence.

SOURCE BF791668.1 GI:12096722

ORGANISM human.

REFERENCE Homo sapiens

AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

TITLE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

CONTACT Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: L1AM962 row: k column: 06  
 High quality sequence stop: 678.  
 Location/Qualifiers

## FEATURES

source

1. 900  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4344317"  
 /clone\_1lb="NIH\_MGC\_84"  
 /tissue\_type="adrenal cortex carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: adrenal gland; Vector: pCMV-SPORT6; Site\_1:  
 NotI; Site\_2: SalI; Cloned unidirectionally; oligo-dT  
 primed. Average insert size 1.229 kb. Library enriched for  
 full-length clones and constructed by Life Technologies.  
 Note: this is a NIH\_MGC Library."

BASE COUNT 283 a 176 c 212 g 229 t

ORIGIN

Query Match 11.8%; Score 657; DB 12; Length 900;

Best Local Similarity 96.5%; Pred. No. 2e-143;  
 Matches 736; Conservative 0; Mismatches 20; Indels 7; Gaps 6;

QY 3356 CTGTATATACAGATTCTCTCCATATCTGATTTTCAGATTAAGACACCTATGGAA 3415  
 Db 1 CTGTATATACAGATTCTCTCCATATCTGATTTTCAGATTAAGACACCTATGGAA 60  
 QY 3416 GTAGTCATGCATCTCAGGTTGTTCTGAGACACCTGATGACTGTTAGATGATGGA 3475  
 Db 61 GTAGTCATGCATCTCAGGTTGTTCTGAGACACCTGATGACTGTTAGATGATGGA 120  
 QY 3476 TAAAGAGATGATGATGTTGCTGAAATGACATTTAGGAAAGTTCGCTGTTTGA 3535  
 Db 121 TAAAGAGATGATGATGTTGCTGAAATGACATTTAGGAAAGTTCGCTGTTTGA 180  
 QY 3536 AAAGCGTCAGAAAGAGAGAGCTTACGAGAGTCCCTACCCCTTACCCATACACATTTGG 3595  
 Db 181 AAAGCGTCAGAGAGAGAGCTTACGAGAGTCCCTACCCCTTACCCATACACATTTGG 240  
 QY 3596 CTGAGGTTACCGAAGAGGGGCCAAGAAATTAGAGTCTTGAAGAGAGACTTATCTAGTG 3655  
 Db 241 CTGAGGTTACCGAAGAGGGGCCAAGAAATTAGAGTCTTGAAGAGAGACTTATCTAGTG 300  
 QY 3656 AGGATGAAGAGCTCCCTGCTCCAAACACTGTTATTTGTTAAAGTAACATATACCTT 3715  
 Db 301 AGGATGAAGAGCTCCCTGCTCCAAACACTGTTATTTGTTAAAGTAACATATACCTT 359  
 QY 3716 CTCAGTCTACTAGGATGACACCGTTGCTACGAGTGTCTGTCTAAGACAGAGAGAGA 3775  
 Db 360 CTCAGTCTACTAGGATGACACCGTTGCTACGAGTGTCTGTCTAAGACAGAGAGAGA 419  
 QY 3776 ATTATTTATGATGAGAAATAGCTTAATGACTGCACTAACAGAGTAATATTTGCCAAAG 3835  
 Db 420 ATTATTTATGATGAGAAATAGCTTAATGACTGCACTAACAGAGTAATATTTGCCAAAG 478  
 QY 3836 CATCTCGAAGAACACCTAGTAGAGAAACAAATGTTCTGCTAGTCTTTCTTCAC 3895  
 Db 479 CATCTCGAAGAACACCTAGTAGAGAAACAAATGTTCTGCTAGTCTTTCTTCAC 538  
 QY 3896 AGTGCATGATTTGAGAGACTTACTGCAATACAAACACCCAGAGATCTTTCTTGATTG 3955  
 Db 539 AGTGCATGATTTGAGAGACTTACTGCAATACAAACACCCAGAGATCTTTCTTGATTG 598  
 QY 3956 GTTCTTCAACAAGATGAGCATCTGTAAGACCGAGGATGTTGCTGATGATGACAGG 4015  
 Db 599 GTTCTTCAACAAGATGAGCATCTGTAAGACCGAGG--TATGCTGCTGATGATGACAGG 656

QY 4016 AATTGCTTACAGATGATGAAGAGAGAGCGGCTTGGAAGAAATATCAAGAGAGC 4075  
 Db 657 AA-TGCTTTACAGATGATGAAGAGAGAG-ACGGCTCGGCAAGAAATATCAAGAGAGC 714  
 QY 4076 AAGCATGATTCAAACTTAGTGAAGCAGCATCTGGGTGTGA 4118  
 Db 715 CAAGCATAG-TTCACACTTAGTGTGACGAGCTGTGTGTGAGA 756

## RESULT 13

LOCUS

BC178466 921 bp mRNA linear EST 06-FEB-2001  
 DEFINITION 60230271F1 NIH\_MGC\_91 Homo sapiens cDNA clone IMAGE:4431568 5',  
 mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

## FEATURES

source

1. 921  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4431568"  
 /clone\_1lb="NIH\_MGC\_91"  
 /tissue\_type="adenocarcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: prostate; Vector: pCMV-SPORT6; Site\_1: NotI;  
 Site\_2: SalI; Cloned unidirectionally; oligo-dT primed.  
 Average insert size 1.4 kb. Library enriched for  
 full-length clones and constructed by Life Technologies.  
 Note: this is a NIH\_MGC Library."

BASE COUNT 340 a 173 c 205 g 203 t

ORIGIN

Query Match 11.1%; Score 619.2; DB 12; Length 921;

Best Local Similarity 90.0%; Pred. No. 1.5e-134;  
 Matches 734; Conservative 0; Mismatches 68; Indels 14; Gaps 6;

QY 2623 CCAAGGAATGACAGAGAGATGTCACATTTCTGCCCCACTGTGGTCTTAAGAAA 2682  
 Db 11 CCAAGGAATGACAGAGAGATGTCACATTTCTGCCCCACTGTGGTCTTAAGAAA 70  
 QY 2683 CAAAGTCCAAAGTCACTTTTGAATGCAACAAAGAAAGAAATCAAGGAAGAAATGAG 2742  
 Db 71 CAAAGTCCAAAGTCACTTTTGAATGCAACAAAGAAAGAAATCAAGGAAGAAATGAG 130  
 QY 2743 TCTAATATCAAGCTGTACAGACAGTAAATATCACTGACAGCTTCTGTGTTGGTCAG 2802  
 Db 131 TCTAATATCAAGCTGTACAGACAGTAAATATCACTGACAGCTTCTGTGTTGGTCAG 190  
 QY 2803 AAAGATTAAGCAGTTGATATGCCAAATGTAGTATCAAGAGGCTCTAGGTTTGTCTA 2862  
 Db 191 AAAGATTAAGCAGTTGATATGCCAAATGTAGTATCAAGAGGCTCTAGGTTTGTCTA 249

OY	2863	ICATCTCAGTTCAGAGGCAAGAAACCTGACATCTACTCCAAATTAACATGAGACTTTTA	2922
Db	250	TCATCTCAGTTCAGAGGCAAGAAACCTGACATCTACTCCAAATTAACATGAGACTTTTA	309
OY	2923	CAAAACCATATCTGATACCAACATTTTTCATCAAGTCAATTTGTATTAACATAAATGT	2982
Db	310	CAAAACCATATCTGATACCAACATTTTTCATCAAGTCAATTTGTATTAACATAAATGT	369
OY	2983	AAGAAAAATCTGCTAGAGAAAACCTTTTGAGAGACATTCATGTCACCTGAAAGAGAAATG	3042
Db	370	AAGAAAAATCTGCTAGAGAAAACCTTTTGAGAGACATTCATGTCACCTGAAAGAGAAATG	429
OY	3043	GGAATAGGAACATTTCCAAAGTACAGTAGAGCAACATTAACCCCTAATTAACATTAGAGAAAT	3102
Db	430	GGAATAGGAACATTTCCAAAGTACAGTAGAGCAACATTAACCCCTAATTAACATTAGAGAAAT	489
OY	3103	GTTTATTAAGAAAGCAGCAGCTCAAGCAATTTTAATGAAGTAGGTCCAGTACTAATGAAGTG	3162
Db	490	GTTTATTAAGAAAGCAGCAGCTCAAGCAATTTTAATGAAGTAGGTCCAGTACTAATGAAGTG	549
OY	3163	GGCTCCAGTATTAATGAATAGGTTCCAGTATGAAACATTCAAGCAGAACTAGTAGA	3222
Db	550	GGCTCCAGTATTAATGAATAGGTTCCAGTATGAAACATTCAAGCAGAACTAGTAGA	609
OY	3223	AACGAGGGCCAAATTAATGATGCTATGCTTAGATTAGGGGTTTTCACACCTGAGG----	3277
Db	610	AACGAGGGGG--AAATTAATGATGCTTAGAATTAAGGGGGTTTTCACAAACCGTAGAGGTC	668
OY	3278	-TCTATAAACCAAGCTCTCTCGAAGTAATGTGAAGCATCTCGTAATTAATAAACCAAGAA	3336
Db	669	TATTAATAAACCAAGGCTCTCTCGGAGTAGATTAAGGC--TCCGGATTAATAAABGGCGAA	726
OY	3337	TATGAGAGAGTAGTTCAGAGCTGTTAATACAGATTTCTCTCCATATCTGATTTTCAGATPAAC	3396
Db	727	ATTGAGAGAGTGTGTCA--ACCGTAACACGATTTTCCCCCATCTCT---GGAATCCGATAAC	782
OY	3397	TTAGAGACGCTTAGGGAAAGTAGATCATCACTCAG	3432
Db	783	TTAAAGCGCCATGGAAAGACCGCAACTAATTGAG	818

```

RESULT 14
LOCUS      BQ215100
DEFINITION BQ215100      878 bp      mRNA      linear      EST 02-MAY-2002
AGENCOURT 7591049 NIH_MGC_72 Homo sapiens CDNA clone IMAGE:6065516
3', mRNA sequence.
ACCESSION  BQ215100
VERSION    BQ215100.1  GI:20396500
KEYWORDS   EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE  1 (bases 1 to 878)
AUTHORS   NIH-MGC hlcp://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.
           Email: cgabs-remail.nih.gov
           Tissue Procurement: ATCC/DCFP/DRP
           CDNA Library Preparation: Life Technologies, Inc.
           DNA Sequencing by: Agencourt Bioscience Corporation
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LM1M1341 row: 0 column: 21
           High quality sequence stop: 669.
           Location/Qualifiers
             1..878
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone="IMAGE:6065516"
FEATURES
SOURCE

```

BASE COUNT	281 a	180 c	190 g	227 t
ORIGIN	/clone_11p="NIH_MGC_72" /tissue_type="melanotic melanoma" /lab_host="DH10B (phage-resistant)" /note="Organ: skin; Vector: pCW-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: oligo dT Average insert size 2 kb. Library constructed by Life Technologies."			

Query Match:	11.0%;	Score	616.2;	DB	14;	Length	878;
Best Local Similarity:	98.7%;	Pred	No. 7.6e-134;				
Matches	664;	Conservative	0;	Mismatches	3;	Indels	6;
						Gaps	4

OY	1	ATGGAATTATCTGGCTTCCTCGCCTTGAAGAAGTACAAAATGTCAATTAAGTCATGCAGAA	60
Db	103	ATGGAATTATCTGGCTTCCTCGCCTTGAAGAAGTACAAAATGTCAATTAAGTCATGCAGAA	162
OY	61	ATCTTAGAGTGCCCATCTGTCTGAGATTGATCAANGAACCTGTCTCACAAAGTGTAC	120
Db	163	ATCTTAGAGTGCCCATCTGTCTGAGATTGATCAANGAACCTGTCTCACAAAGTGTAC	222
OY	121	CACATATTTTGGCAAATTTTGGCATGCTGAAACTCTCAACCAGAAAGAGGCCCTTACAG	180
Db	223	CACATATTTTGGCAAATTTTGGCATGCTGAAACTCTCAACCAGAAAGAGGCCCTTACAG	282
OY	181	TGTCTCTTATGTAAAGATGATATAACAAGAGAGAGCCCTACAAAGATACGATTTAAGT	240
Db	283	TGTCTCTTATGTAAAGATGATATAACAAGAGAGAGCCCTACAAAGATACGATTTAAGT	342
OY	241	CAACTTGTTSAGAGCATATTGAAATCAATTTGTGCTTTTCAAGCTTGCACAGTITGGAG	300
Db	343	CAACTTGTTSAGAGCATATTGAAATCAATTTGTGCTTTTCAAGCTTGCACAGTITGGAG	402
OY	301	TATGCAACAGCATATATTTTTCCAAGAAAAAGAAATAACTCTCCTGAACATPAAAGAT	360
Db --	403	TATGCAACAGCATATATTTTTCCAAGAAAAAGAAATAACTCTCCTGAACATPAAAGAT	462
OY	361	GAAATTTCATCATCCAAAGTATGGGCTACAGAAACCGTGGCCAAAGACATTCTACAGAT	420
Db	463	GAAATTTCATCATCCAAAGTATGGGCTACAGAAACCGTGGCCAAAGACATTCTACAGAT	522
OY	421	GAACCCGAAATCTCTTCCTTGCAGAAACAGTCTAGTGTCCAACTCTTAACCTTGA	480
Db	523	GAACCCGAAATCTCTTCCTTGCAGAAACAGTCTAGTGTCCAACTCTTAACCTTGA	579
OY	481	ACGTGTGAGAACTGTGAGCAAGAGAGGSGATACACCTCAAAAAGAGCTGTCTACAT	540
Db	580	ACGTGTGAGAACTGTGAGCAAGAGAGGSGATACACCTCAAAAAGAGCTGTCTACAT	639
OY	541	GAATTGGAGATCGATTCCTTGAAGATACCGTTAATAAGCAACTTATTTGCACT -GTGG	599
Db	640	GAATTGGAGATCGATTCCTTGAAGATACCGTTAATAAGCAACTTATTTGCACTGTGG	699
OY	600	AGATCAAGATTTGTTACAATACACCCCTCAAGAACCA -GGGATGAATCAG -TTTGGAT	657
Db	700	AGATCAAGATTTGTTACAATACACCCCTCAAGAACCAAGGAGATGAATCAGTTTGGAT	759
OY	658	TCTGCAAAAAGG 670	
Db	760	TCTGCAAAAAGG 772	

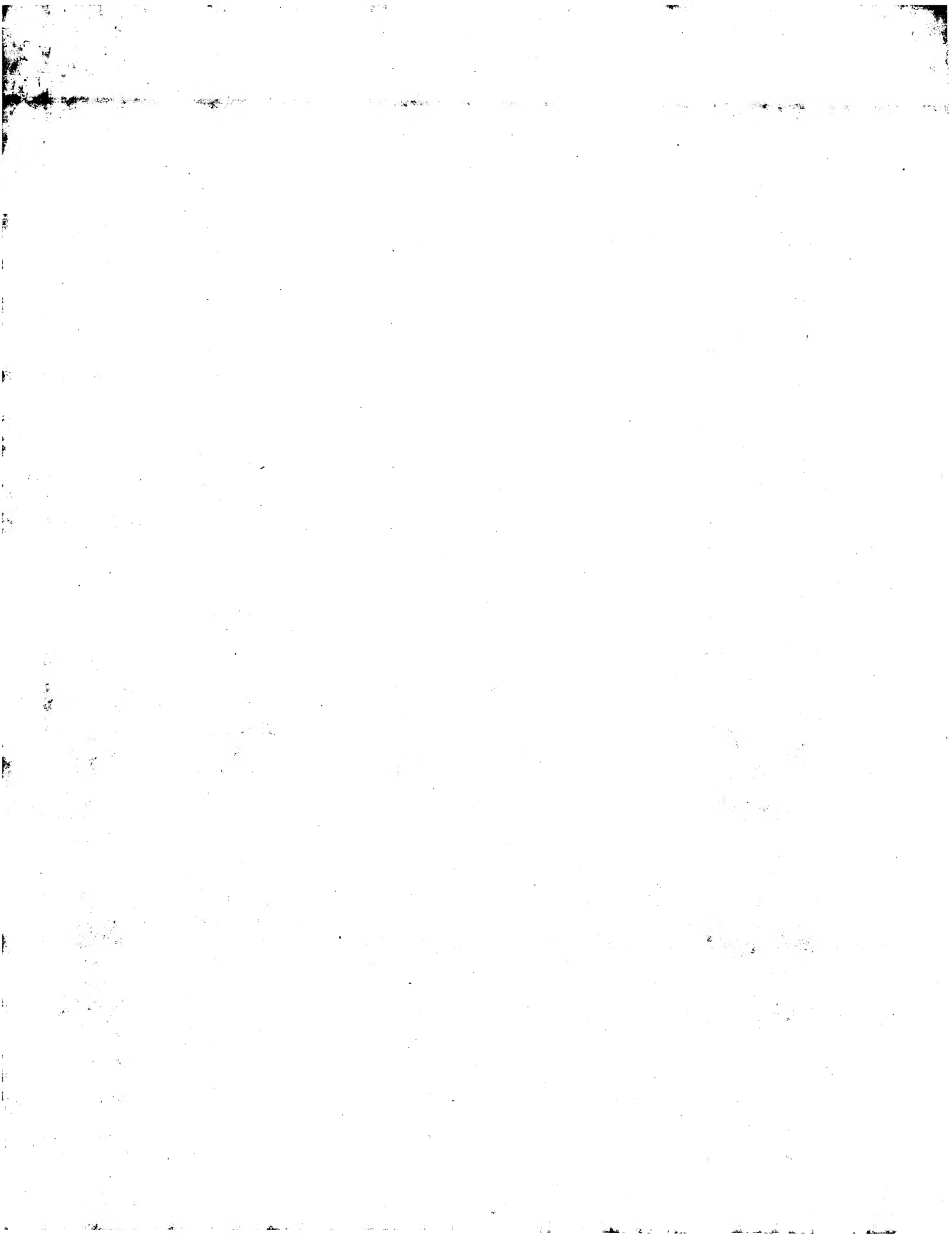
RESULT 15	
A1992040/c	
LOCUS	A1992040 630 bp mRNA linear EST 08-SEP-1999
DEFINITION	ws44c12.x1 NCI-GAP_Brn25 Homo sapiens cDNA clone IMAGE:250054 3'
	(HUMAN); mRNA sequence.
ACCESSION	A1992040
VERSION	A1992040.1 GI:5838945
KEYWORDS	EST.
SOURCE	human.

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 630)  
 AUTHORS NCI/NINDS-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute / National Institute of Neurological  
 Disorders and Stroke, Brain Tumor Genome Anatomy Project  
 (CGAP/BTGP), Tumor Gene Index  
 Unpublished (1998)  
 JOURNAL Contact: Robert Strausberg, Ph.D.  
 COMMENT Email: cgaps@emall.nih.gov  
 Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,  
 Ph.D.  
 CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
 Bonaldo, Ph.D.  
 CDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 www.bio.lnl.gov/bdrrp/image/image.html  
 Seq primer: -40up from Gibco  
 High quality sequence stop: 453.  
 FEATURES Location/Qualifiers  
 Source 1. 630  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2500054"  
 /clone\_1lb="NCI-CGAP\_Brn25"  
 /tissue\_type="anaplastic oligodendroglioma"  
 /lab\_host="DH10B"  
 /note="Organ: brain; Vector: pT73D-Pac (Pharmacia) with a  
 modified polylinker; Site: 1: Not I; Site: 2: Eco RI; 1st  
 strand cDNA was primed with a Not I - oligo(dT) primer [5'  
 TGTTCACCAATCTGAGTGAAGGAGCGGCGCATAGTTTCTTTTCTTTTCTTTT  
 T 3']; double-stranded cDNA was ligated to Eco RI  
 adaptors (Pharmacia), digested with Not I and cloned into  
 the Not I and Eco RI sites of the modified pT73 vector.  
 Library is normalized, and was constructed by Bento  
 Soares and M. Fatima Bonaldo."  
 BASE COUNT 146 a 118 c 111 g 252 t 3 others  
 ORIGIN  
 Query Match 10.8%; Score 603.8; DB 9; Length 630;  
 Best Local Similarity 98.3%; Pred. No. 5.6e-131;  
 Matches 619; Conservative 0; Mismatches 10; Indels 1; Gaps 1;  
 QY 1321 ATATGTAAAGTGAAGAGTTCCTCAATCAGTAGAGATATTTGAAGACAAATA 1380  
 |||||||  
 DB 630 ATATGTAAAGTGAAGAGTTCCTCAATCAGTAGAGATATTTGAAGACAAATA 571  
 QY 1381 TTGGGAAAACCTATCGAAGAGGCAAGGC-TCCCCAATTAAAGCATGTAAGTGA 1439  
 |||||||  
 DB 570 TTGGGAAAACCTATCGAAGAGGCAAGGC-TCCCCAATTAAAGCATGTAAGTGA 511  
 QY 1440 TCTAATTATAGACATTTGTTACTAGCCACAGATATATCAAGAGCGTCCCTCACAAA 1499  
 |||||||  
 DB 510 TCTAATTATAGACATTTGTTACTAGCCACAGATATATCAAGAGCGTCCCTCACAAA 451  
 QY 1500 TAAATTAAGCGTAAAGAGACCTACATAGCGCTTCATCTGAGGATTTTATCAAGAA 1559  
 |||||||  
 DB 450 TAAATTAAGCGTAAAGAGACCTACATAGCGCTTCATCTGAGGATTTTATCAAGAA 391  
 QY 1560 AGCAGATTTGGCGTTCAAAAAGACTCTGAAATGATTAATCAGGAACTAACCAAGCGA 1619  
 |||||||  
 DB 390 AGCAGATTTGGCGTTCAAAAAGACTCTGAAATGATTAATCAGGAACTAACCAAGCGA 331  
 QY 1620 GCAGAAATGCTCAAGTGAATGATTTAATAAGTGTGTCATGAGATTAACCAAGAGTGA 1679  
 |||||||  
 DB 330 GCAGAAATGCTCAAGTGAATGATTTAATAAGTGTGTCATGAGATTAACCAAGAGTGA 271  
 QY 1680 TTCTATTCAAGATGAGAAAAATCCTTAACCAATAGAAATCACTCGAAAAAGAAATCTGCTTT 1739  
 |||||||  
 DB 270 TTCTATTCAAGATGAGAAAAATCCTTAACCAATAGAAATCACTCGAAAAAGAAATCTGCTTT 211

QY 1740 CAAACGAAAGCTGACCTATTAAGCAGATATAGCAATATGGAACCTCGAATTAATAT 1799  
 |||||||  
 DB 210 CAAACGAAAGCTGACCTATTAAGCAGATATAGCAATATGGAACCTCGAATTAATAT 151  
 QY 1800 CCACAAATTCAAAGACACCTAAAGAAATAGGCTGAGAGGAAAGTCTTCTACAGGATAT 1859  
 |||||||  
 DB 150 CCACAAATTCAAAGACACCTAAAGAAATAGGCTGAGAGGAAAGTCTTCTACAGGATAT 91  
 QY 1860 TCATGCGCTTGAAGTACTAGTACAGTACAGAAATCTTAAGCCACCTATTGTACTGAATTGCA 1919  
 |||||||  
 DB 90 TCATGCGCTTGAAGTACTAGTACAGTACAGAAATCTTAAGCCACCTATTGTACTGAATTGCA 31  
 QY 1920 AATTGATAGTTGTTCTGACAGTGAAGAT 1949  
 |||||||  
 DB 30 AATTGATAGTTGTTCTGACAGTGAAGAT 1

Search completed: June 27, 2003, 05:00:08  
 Job time : 7063.55 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: June 26, 2003, 12:34:02 ; Search time 13911.7 Seconds

(Without alignments)  
11692.029 Million cell updates/sec

Title: US-09-734-672-3\_COPY\_120\_5708

Perfect score: 5589

Sequence: 1 ATGATTTATCTGCTCTTCG.....AGATCCCCACAGCCACTAC 5589

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapept 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

GenEmbl:\*

- 1: gb\_da:\*
- 2: gb\_htg:\*
- 3: gb\_in:\*
- 4: gb\_ov:\*
- 5: gb\_ov:\*
- 6: gb\_pat:\*
- 7: gb\_ph:\*
- 8: gb\_pl:\*
- 9: gb\_pr:\*
- 10: gb\_ro:\*
- 11: gb\_sts:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: gb\_vi:\*
- 15: em\_ba:\*
- 16: em\_fun:\*
- 17: em\_hum:\*
- 18: em\_in:\*
- 19: em\_mu:\*
- 20: em\_om:\*
- 21: em\_or:\*
- 22: em\_ov:\*
- 23: em\_pat:\*
- 24: em\_ph:\*
- 25: em\_pl:\*
- 26: em\_ro:\*
- 27: em\_sts:\*
- 28: em\_un:\*
- 29: em\_vi:\*
- 30: em\_htg\_hum:\*
- 31: em\_htg\_inv:\*
- 32: em\_htg\_other:\*
- 33: em\_htg\_mus:\*
- 34: em\_htg\_pln:\*
- 35: em\_htg\_rtd:\*
- 36: em\_htg\_man:\*
- 37: em\_htg\_vrt:\*
- 38: em\_sy:\*
- 39: em\_htgo\_hum:\*
- 40: em\_htgo\_mus:\*
- 41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5587.4	100.0	5711	6 AR007334	AR007334 Sequence
2	5587.4	100.0	5711	6 AR112810	AR112810 Sequence
3	5585.8	99.9	5711	6 AR033056	AR033056 Sequence
4	5585.8	99.9	5711	6 HS014680	U14680 Homo sapien
5	5585.8	99.9	5712	6 AR070223	AR070223 Sequence
6	5585.8	99.9	5712	6 AR118507	AR118507 Sequence
7	5585.8	99.9	5712	6 AR125601	AR125601 Sequence
8	5585.8	99.9	5712	6 AR184044	AR184044 Sequence
9	5585.8	99.9	5914	6 AR004673	AR004673 Sequence
10	5585.8	99.9	5914	6 AR008159	AR008159 Sequence
11	5585.8	99.9	5914	6 AR136942	AR136942 Sequence
12	5585.8	99.9	5914	6 I76943	I76943 Sequence 1
13	5585.8	99.9	5914	6 I80938	I80938 Sequence 1
14	5585.8	99.9	5914	6 I81034	I81034 Sequence 1
15	5584.2	99.9	5711	6 AR048660	AR048660 Sequence
16	5584.2	99.9	5711	6 AR048666	AR048666 Sequence
17	5584.2	99.9	5711	6 I40795	I40795 Sequence 4
18	5584.2	99.9	5711	6 I40801	I40801 Sequence 10
19	5582.6	99.9	5711	6 AR007335	AR007335 Sequence
20	5582.6	99.9	5711	6 AR112809	AR112809 Sequence
21	5581	99.9	5711	6 AR007333	AR007333 Sequence
22	5581	99.9	5711	6 AR112808	AR112808 Sequence
23	5581	99.9	5711	6 I59546	I59546 Sequence 1
24	5574.8	99.7	5712	6 AR048668	AR048668 Sequence
25	5574.8	99.7	5712	6 I40803	I40803 Sequence 12
26	5573.8	99.7	5710	6 AR048662	AR048662 Sequence
27	5573.8	99.7	5710	6 I40797	I40797 Sequence 6
28	5571.8	99.7	5709	6 AR048658	AR048658 Sequence
29	5571.8	99.7	5709	6 AR048663	AR048663 Sequence
30	5571.8	99.7	5709	6 AR048664	AR048664 Sequence
31	5571.8	99.7	5709	6 AR048665	AR048665 Sequence
32	5571.8	99.7	5709	6 I40793	I40793 Sequence 2
33	5571.8	99.7	5709	6 I40798	I40798 Sequence 7
34	5571.8	99.7	5709	6 I40799	I40799 Sequence 8
35	5571.8	99.7	5709	6 I40800	I40800 Sequence 9
36	5567.8	99.6	5707	6 AR048667	AR048667 Sequence
37	5567.8	99.6	5707	6 I40802	I40802 Sequence 11
38	5535.8	99.0	5709	6 AX281859	AX281859 Sequence
39	5531.8	99.0	5689	6 AR048659	AR048659 Sequence
40	5531.8	99.0	5689	6 I40794	I40794 Sequence 3
41	5526.6	98.9	5589	9 AF207822	AF207822 Pan trogl
42	5507.8	98.5	5693	9 AF005068	AF005068 Homo sapi
43	5505.6	98.5	5770	6 AR048661	AR048661 Sequence
44	5505.6	98.5	5770	6 I40796	I40796 Sequence 5
45	5467.8	97.8	5656	6 AR048657	AR048657 Sequence

## ALIGNMENTS

RESULT 1  
AR007334  
LOCUS AR007334 5711 bp DNA  
DEFINITION Sequence 3 from patent US 5750400. Linear PAT 04-DEC-1998  
ACCESSION AR007334  
VERSION AR007334.1 GI:3966818  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 5711)  
AUTHORS Murphy,P.D., Allen,A.C., Alvares,C.P., Critz,B.S., Olson,S.J.,  
Schelter,D.B. and Zeng,B.  
TITLE Coding sequences of the human BRCA1 gene  
JOURNAL Patent: US 5750400-A 3 12-MAY-1998;

FEATURES Location/Qualifiers  
source 1..5711  
/organism="unknown"  
BASE COUNT 1956 a 1098 c 1274 g 1383 t  
ORIGIN  
Query Match 100.0%; Score 5587.4; DB 6; Length 5711;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 5588; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
1 AAGGATTTTCGCTTCCTGGCCCTGGAAGAGTCAAAATGCTCTTAATGCTATGCGAAA 60  
120 AAGGATTTTCGCTTCCTGGCCCTGGAAGAGTCAAAATGCTCTTAATGCTATGCGAAA 179  
61 ATCTTAGAGTGTCCATCTGCTGAGAGTGTATCAAGAACCTGTCTCCACAAAGTGTAC 120  
180 ATCTTAGAGTGTCCATCTGCTGAGAGTGTATCAAGAACCTGTCTCCACAAAGTGTAC 239  
121 CACATATTTTGCATAATTTTGCATGCTGAACCTTCTCAACCCAGAAAGGCCCTTACAG 180  
240 CACATATTTTGCATAATTTTGCATGCTGAACCTTCTCAACCCAGAAAGGCCCTTACAG 299  
181 TGTCCCTTTATGTAAGATATATATACCAAAAGAGCCCTACAGAAAGTACAGATTTAGT 240  
300 TGTCCCTTTATGTAAGATATATATACCAAAAGAGCCCTACAGAAAGTACAGATTTAGT 359  
241 CAACCTTGTGAAGAGCTATTTGAATAATCATTTTGTCTTTCAGCTTGCACAGAGTTTGAG 300  
360 CAACCTTGTGAAGAGCTATTTGAATAATCATTTTGTCTTTCAGCTTGCACAGAGTTTGAG 419  
301 TATGCAACAGCTATATATTTTGCATAAAAGGAAATTAACCTCTCTGACATCTTAANAAT 360  
420 TATGCAACAGCTATATATTTTGCATAAAAGGAAATTAACCTCTCTGACATCTTAANAAT 479  
361 GAAGTTTTCATATCCAAAGTATGAGCTACAGAAACCGTCCAAAGAGCTTACAGAGT 420  
480 GAAGTTTTCATATCCAAAGTATGAGCTACAGAAACCGTCCAAAGAGCTTACAGAGT 539  
421 GAACCCGAAATTCCTCTCTTGCAGGAAACCAAGCTTCTAGTCCAACTCTTAACCTTTGA 480  
540 GAACCCGAAATTCCTCTCTTGCAGGAAACCAAGCTTCTAGTCCAACTCTTAACCTTTGA 599  
481 ACTGTGAGAACTCTGAGCAAGCAAGCGGGATACAACTCTCAAAAGAGCTGTCTTAATTT 540  
600 ACTGTGAGAACTCTGAGCAAGCAAGCGGGATACAACTCTCAAAAGAGCTGTCTTAATTT 659  
541 GAATGGGATCGATTTCTCTGAGATACCGTTTAATGAAGCAACTATGCAAGTGGGA 600  
660 GAATGGGATCGATTTCTCTGAGATACCGTTTAATGAAGCAACTATGCAAGTGGGA 719  
601 GATCAGAATTTGTTCAAAATCAACCCCTCAAGSAGACAGGATGAATCAGTTTGGATTCT 660  
720 GATCAGAATTTGTTCAAAATCAACCCCTCAAGSAGACAGGATGAATCAGTTTGGATTCT 779  
661 GCAAAAAGGCTGCTTGTGAATTTTCTGAGCGGATGTAAACAATTAAGATCAATCA 720  
780 GCAAAAAGGCTGCTTGTGAATTTTCTGAGCGGATGTAAACAATTAAGATCAATCA 839  
721 CCCAGTAAATGATTTTGAACACCACTGAGAGCGGAGCTGAGAGGATCCAGAAAG 780  
840 CCCAGTAAATGATTTTGAACACCACTGAGAGCGGAGCTGAGAGGATCCAGAAAG 899  
781 TATCAGGGTAGTCTGTTCAAACTTGCATGTGAGCCATGTGGCAACAATACTCTATGCC 840  
900 TATCAGGGTAGTCTGTTCAAACTTGCATGTGAGCCATGTGGCAACAATACTCTATGCC 959  
841 AGCTCATTAACAGATGAGACAGCAAGTTTATTACTACTTAAGACAGATGAATGTAGA 900  
960 AGCTCATTAACAGATGAGACAGCAAGTTTATTACTACTTAAGACAGATGAATGTAGA 1019  
901 AAGGCTGAATTTCTGAATTAAGCAAAACAGCTGCTTGAAGAGCAACATATAACAA 960  
1020 AAGGCTGAATTTCTGAATTAAGCAAAACAGCTGCTTGAAGAGCAACATATAACAA 1079

QY 961 TGGGCTGGAAGTAAAGCAACATGTATGATTAAGCGGAGCTCCGACAGCAAGAAAAAGGTA 1020  
DB 1080 TGGGCTGGAAGTAAAGCAACATGTATGATTAAGCGGAGCTCCGACAGCAAGAAAAAGGTA 1139  
QY 1021 GATCTGAATGCTGATCCCTGTGTGAGAGAAAAAGATGAATTAAGCAGAAATCGCATGC 1080  
DB 1140 GATCTGAATGCTGATCCCTGTGTGAGAGAAAAAGATGAATTAAGCAGAAATCGCATGC 1199  
QY 1081 TCAGAAATTCCTAGAGATCTGAGAGTCTCTTGATTAACACTTAATTAAGCAGATTCAG 1140  
DB 1200 TCAGAAATTCCTAGAGATCTGAGAGTCTCTTGATTAACACTTAATTAAGCAGATTCAG 1259  
QY 1141 AAGGTTAATGAGTGTCTTTCAGAACTGATGAGTGTAGTCTGATGAGTCAATGAT 1200  
DB 1260 AAGGTTAATGAGTGTCTTTCAGAACTGATGAGTGTAGTCTGATGAGTCAATGAT 1319  
QY 1201 GGGGAGTCTGAATCAAAATGCCAAAGTACGTATTTGAGAGCTTCAATTAATGAGTACAT 1260  
DB 1320 GGGGAGTCTGAATCAAAATGCCAAAGTACGTATTTGAGAGCTTCAATTAATGAGTACAT 1379  
QY 1261 GAATATTCGTGTTCTTCAAGAAATTAAGTACCTAGGCGAGTACCTCATGAGGCTTTA 1320  
DB 1380 GAATATTCGTGTTCTTCAAGAAATTAAGTACCTAGGCGAGTACCTCATGAGGCTTTA 1439  
QY 1321 ATATGTAAAGTCAAAAGAGTTCATCTCAAAATCAGTAGAGAGTAAATTTGAAGACAAATA 1380  
DB 1440 ATATGTAAAGTCAAAAGAGTTCATCTCAAAATCAGTAGAGAGTAAATTTGAAGACAAATA 1499  
QY 1381 TTTGGGAAACCTATTCGAGAGAGCAAGCCCTCCCAACTTAAGCATGTAACGAAAT 1440  
DB 1500 TTTGGGAAACCTATTCGAGAGAGCAAGCCCTCCCAACTTAAGCATGTAACGAAAT 1559  
QY 1441 CTAAATTAAGAGCAATTTGTTACTGAGCCAGCATATTAACAAAGAGCCCTCCCAACT 1500  
DB 1560 CTAAATTAAGAGCAATTTGTTACTGAGCCAGCATATTAACAAAGAGCCCTCCCAACT 1619  
QY 1501 AAATTAAGCGTAAAGAGAGACCTACATCAGGCTTCACTCGAGAGATTTTATCAAGAA 1560  
DB 1620 AAATTAAGCGTAAAGAGAGACCTACATCAGGCTTCACTCGAGAGATTTTATCAAGAA 1679  
QY 1561 GCAGATTTGGCAGTTCAAAAGACCTCTGTAATGATTAATCAAGGAACTAACCAAGGAG 1620  
DB 1680 GCAGATTTGGCAGTTCAAAAGACCTCTGTAATGATTAATCAAGGAACTAACCAAGGAG 1739  
QY 1621 CAGAAATGCTCAAGTATGATTAATTAATAGTGTCTGAGAAATTAACAAAGAGGAT 1680  
DB 1740 CAGAAATGCTCAAGTATGATTAATTAATAGTGTCTGAGAAATTAACAAAGAGGAT 1799  
QY 1681 TCTATTCAGAAATGAGAAAAATCCTAACCCATTAAGATCACTGAAAAAGAAATTCGCTTTC 1740  
DB 1800 TCTATTCAGAAATGAGAAAAATCCTAACCCATTAAGATCACTGAAAAAGAAATTCGCTTTC 1859  
QY 1741 AAAAGCAAAAGCTGAACCTATTAAGACAGCAAGTATTAAGCAATTAAGCAATTCGAATTAATATC 1800  
DB 1860 AAAAGCAAAAGCTGAACCTATTAAGACAGCAAGTATTAAGCAATTAAGCAATTCGAATTAATATC 1919  
QY 1801 CACAATTCAAAAGCAACCTTAAGAAAGATAGGCGAGAGAGAGTCTTACACAGGCATATT 1860  
DB 1920 CACAATTCAAAAGCAACCTTAAGAAAGATAGGCGAGAGAGAGTCTTACACAGGCATATT 1979  
QY 1861 CATGGCCTTGAAGTATGATGATGATGATTAAGCCCACTTAATTTAGTGAATTTGCAA 1920  
DB 1980 CATGGCCTTGAAGTATGATGATGATGATTAAGCCCACTTAATTTAGTGAATTTGCAA 2039  
QY 1921 ATTGATAGTGTCTTACAGAGTGAAGAGATTAAGAAAAAGTCAACAAATTCAGTC 1980  
DB 2040 ATTGATAGTGTCTTACAGAGTGAAGAGATTAAGAAAAAGTCAACAAATTCAGTC 2099  
QY 1981 AGGCACAGAGAAACCTCAACTCATGAGAGTAAAGAACCTGAGAGCCCAAGAAAG 2040  
DB 2100 AGGCACAGAGAAACCTCAACTCATGAGAGTAAAGAACCTGAGAGCCCAAGAAAG 2159

Db	3240	TCAACCAATATTAATGAAGTAGGGTTCACGTAAGTAAAGAGGGGCTCCAGATTATATGAA	3299
OY	3181	ATAGGTTCACATGATGAAGAAATTCACACACAGACTAGCTAAGAAACAGAGGGCCAAAATTG	3240
Db	3300	ATAGGTTCACATGATGAAGAAATTCACACAGACTAGCTAAGAAACAGAGGGCCAAAATTG	3359
OY	3241	AATGCTATGCTTGAATTAAGGGGTTTGGCAACCTGAGGCTATTAACAAAGCTTCTCGGA	3300
Db	3360	AATGCTATGCTTGAATTAAGGGGTTTGGCAACCTGAGGCTATTAACAAAGCTTCTCGGA	3419
OY	3301	AGTATTTTAAGCATCCTGGAATTAAGAAAGCAAGAAATTTGAAGAAAGTATGTTCAACACTG	3360
Db	3420	AGTATTTTAAGCATCCTGGAATTAAGAAAGCAAGAAATTTGAAGAAAGTATGTTCAACACTG	3479
OY	3361	AATACAGATTTCTGCTCCATATCTGATTTACATTAAGTAAAGAGCCATGGAAGTAGT	3420
Db	3480	AATACAGATTTCTGCTCCATATCTGATTTACATTAAGTAAAGAGCCATGGAAGTAGT	3539
OY	3421	CATGCATCTCAGGTTTGTCTTGAGACACCTATGACCTGTTAGATGATGTAATTAAG	3480
Db	3540	CATGCATCTCAGGTTTGTCTTGAGACACCTATGACCTGTTAGATGATGTAATTAAG	3599
OY	3481	GAGATCTACTGTTTGGCTGAAATATGACATTAAAGAAAGTTCTGCTGTTTAAAGAAAGC	3540
Db	3600	GAGATCTACTGTTTGGCTGAAATATGACATTAAAGAAAGTTCTGCTGTTTAAAGAAAGC	3659
OY	3541	GTCCGAAAGAGAGGCTTAGCAGAGAGTCTAGCCCTTTCACCCATATACATTTGGCTCAG	3600
Db	3660	GTCCGAAAGAGAGGCTTAGCAGAGAGTCTAGCCCTTTCACCCATATACATTTGGCTCAG	3719
OY	3601	GGTTACCCAGAGAGGGGCCAAGAAATTAAGTCTCAGAGAGAACTTATCTAGTACAGAT	3660
Db	3720	GGTTACCCAGAGAGGGGCCAAGAAATTAAGTCTCAGAGAGAACTTATCTAGTACAGAT	3779
OY	3661	GAAAGGCTTCCCTGCTTCCAAACACTGTTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG	3720
Db	3780	GAAAGGCTTCCCTGCTTCCAAACACTGTTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG	3839
OY	3721	TCTCTAGCAGTACGACCGTGTGACGAGAGTGTCTAGTAAAGACACAGAGAAATTTA	3780
Db	3840	TCTCTAGCAGTACGACCGTGTGACGAGAGTGTCTAGTAAAGACACAGAGAAATTTA	3899
OY	3781	TTATCATTTGAAGAAATAGCTTAAATGACATGCAATTAACCGATTAATTTGGCAAGGCAAT	3840
Db	3900	TTATCATTTGAAGAAATAGCTTAAATGACATGCAATTAACCGATTAATTTGGCAAGGCAAT	3959
OY	3841	CAGGAACATCACCTTAGTGAGAGAAACAAATGTTCTGCTAGCTGTTTCTTCAACAGTGC	3900
Db	3960	CAGGAACATCACCTTAGTGAGAGAAACAAATGTTCTGCTAGCTGTTTCTTCAACAGTGC	4019
OY	3901	AGTGAATTTGGAAGAGCTGACGTGCAAAATTCAAACCCAGAGATTCCTTTCTGATTTGGTCT	3960
Db	4020	AGTGAATTTGGAAGAGCTGACGTGCAAAATTCAAACCCAGAGATTCCTTTCTGATTTGGTCT	4079
OY	3961	TCCAAAACAAATGAGGCACTACGTGCAAAAGCCAGAGGAGTTGGTCTGAGTGAACAGGAATTG	4020
Db	4080	TCCAAAACAAATGAGGCACTACGTGCAAAAGCCAGAGGAGTTGGTCTGAGTGAACAGGAATTG	4139
OY	4021	GTTCACAGTGTGAAGAAAGAGAAACGGGCTTGGAGAGAAATTAATCAAGAGACCAAGC	4080
Db	4140	GTTCACAGTGTGAAGAAAGAGAAACGGGCTTGGAGAGAAATTAATCAAGAGACCAAGC	4199
OY	4081	ATGGAATTCAAACTTAGGTGAAGAGAGCAATCTGGGGTGAAGAGTGAACAAAGCTCTCGAA	4140
Db	4200	ATGGAATTCAAACTTAGGTGAAGAGAGCAATCTGGGGTGAAGAGTGAACAAAGCTCTCGAA	4259
OY	4141	GACTGCTCAGAGGCTATCCTCTCAAGAGTACATTTTAACACTCAGACAGAGGATATCATG	4200
Db	4260	GACTGCTCAGAGGCTATCCTCTCAAGAGTACATTTTAACACTCAGACAGAGGATATCATG	4319
OY	4201	CACATTAACCTGATTAAGCTCAGCAGAGAAATGCTGAATGAAGCTGTGTGAACAG	4260

Dp	4320	CAACATTAACCTGATTAAGGCTCCAGCAGGAAATGGCTGAACCTAAGACGTGTGTAGAACAG	4379
Qy	4261	CATGGAGGCGACGCTTCTTACAGCTACCCCTTCATCATTAAGTACTCTTCTGCCTTGAG	4320
Dp	4380	CATGGGAGGCGACGCTTCTTACAGCTACCCCTTCATCATTAAGGACTCTTCTGCCTTGAG	4439
Qy	4321	GACCTGGCGAAATCCGAAGCAAAAGCAGATCAGAAAAAGAGATTAACTTCCACGAAAGCT	4380
Dp	4440	GACCTGGCGAAATCCGAAGCAAAAGCAGATCAGAAAAAGAGATTAACTTCCACGAAAGT	4499
Qy	4381	AGTGAATACCCCTTAAGCCAGAAATCCAGAAAGGCGCTTTCGTGACAACTTTGAGGTGCT	4440
Dp	4500	AGTGAATACCCCTTAAGCCAGAAATCCAGAAAGGCGCTTTCGTGACAACTTTGAGGTGCT	4559
Qy	4441	GCAGATAGTCTTCCACAGTAAAAATTAAGACCGAGAGTGGAAAGTCACTCCCTTTTAA	4500
Dp	4560	GCAGATAGTCTTCCACAGTAAAAATTAAGACCGAGAGTGGAAAGTCACTCCCTTTTAA	4619
Qy	4501	TGCCCATCATTTAGATGATAGGAGTGCATGACATTCCTCGGGAGCTTCAGATTGA	4560
Dp	4620	TGCCCATCATTTAGATGATAGGAGTGCATGACATTCCTCGGGAGCTTCAGATTGA	4679
Qy	4561	AACTAACCATCTCAAGAGAGGCTCATTTAAGTTGTTGATGTGAGAGCAACAGCTGGAA	4620
Dp	4680	AACTAACCATCTCAAGAGAGGCTCATTTAAGTTGTTGATGTGAGAGCAACAGCTGGAA	4739
Qy	4621	GAGTGTGGGCCACACGATTTTGAACGAAACATCTTACTTCCAAAGCAAGATCTAGAGGA	4680
Dp	4740	GAGTGTGGGCCACACGATTTTGAACGAAACATCTTACTTCCAAAGCAAGATCTAGAGGA	4799
Qy	4681	ACCCCTTACCTGGAACTGAGATCAAGCCTCTTCTGTGATGACCCGAACTTGATCTTCT	4740
Dp	4800	ACCCCTTACCTGGAACTGAGATCAAGCCTCTTCTGTGATGACCCGAACTTGATCTTCT	4859
Qy	4741	GAAGACAGAGCCCGCAGAGTCAAGCTCTTGGCAACATACCATCTTCAACTCTGCATTG	4800
Dp	4860	GAAGACAGAGCCCGCAGAGTCAAGCTCTTGGCAACATACCATCTTCAACTCTGCATTG	4919
Qy	4801	AAAGTTCCTCCCAATTTGAAAGTTGCAGAAATCTGCCAGAGTCCAGCTGCTGCATACTACT	4860
Dp	4920	AAAGTTCCTCCCAATTTGAAAGTTGCAGAAATCTGCCAGAGTCCAGCTGCTGCATACTACT	4979
Qy	4861	GATACCTGCTGGGTAAATGCAATGGAACAAAGTGTGACAGGAGGAAGCCAGATTGACA	4920
Dp	4980	GATACCTGCTGGGTAAATGCAATGGAACAAAGTGTGACAGGAGGAAGCCAGATTGACA	5039
Qy	4921	GCTTCAACAGAAAGGCTCAACAAAAGATGTCCATGCTGTGTGTGGCTTACCCCGAA	4980
Dp	5040	GCTTCAACAGAAAGGCTCAACAAAAGATGTCCATGCTGTGTGTGGCTTACCCCGAA	5099
Qy	4981	GAATTTATGCTGTGTACAGTTCGCCAGAAACCAACATCACTTTAACTAATCTAATT	5040
Dp	5100	GAATTTATGCTGTGTACAGTTCGCCAGAAACCAACATCACTTTAACTAATCTAATT	5159
Qy	5041	ACTGAAGAGCACTACTCATGTGTTGATGAAAACACATGCTGATGTGTGTGAACGACA	5100
Dp	5160	ACTGAAGAGCACTACTCATGTGTTGATGAAAACACATGCTGATGTGTGTGAACGACA	5219
Qy	5101	CTGAAGATTTTCTAGCAATTGGCGAGAGAAATGGAGTACTCTATTTCTGGGTGACC	5160
Dp	5220	CTGAAGATTTTCTAGCAATTGGCGAGAGAAATGGAGTACTCTATTTCTGGGTGACC	5279
Qy	5161	CAGTCTATTAAAGAAGAAATCTGTGAATGAGATGATTTTGAAGTCCAGAGAGATGTG	5220
Dp	5280	CAGTCTATTAAAGAAGAAATCTGTGAATGAGATGATTTTGAAGTCCAGAGAGATGTG	5339
Qy	5221	GTCAATGAGAAACACCAAGGTCCAAAGCGACCAAGAGAAATCCACAGACAGAAAGTC	5280
Dp	5340	GTCAATGAGAAACACCAAGGTCCAAAGCGACCAAGAGAAATCCACAGACAGAAAGTC	5399
Qy	5281	TTCAAGGGGGCTAGAAATCTGTTGTTATGGGCGCTTACCAACATGCGCCACAGATCAACTG	5340
Dp	5400	TTCAAGGGGGCTAGAAATCTGTTGTTATGGGCGCTTACCAACATGCGCCACAGATCAACTG	5459

QY		5341	GATGAGTGTACAGCTGTGTGTCTCCTTCGTGTGGTAAGAGACTTTCAATTCACCCCT	5400
Db		5460	GATGCGATGTACAAGCTGTGTGTCTTCTTGCTGTGAAGAAGACTTTCATCTCACCCCT	5519
QY		5401	GGCAGAGTGTGCACCACAATTGTGTGGTTGTGCAGCCAGATGCCCTGGAGACAGAAATGC	5460
Db		5520	GGCAGAGTGTGCACCACAATTGTGTGGTTGTGCAGCCAGATGCCCTGGAGACAGAAATGC	5579
QY		5461	TTCATGTCAATTGGGCGACATGTGTGTAGGCACCTGTGTGTACCCGAGAGTGGGTGTGGAC	5520
Db		5580	TTCATGTCAATTGGGCGAGATGTGTGTAGGCACCTGTGTGTACCCGAGAGTGGGTGTGGAC	5639
QY		5521	AGTGTAGCACTCTTACAGTGCAGGAGCGAGAGCTGGACACTCTCGTAGTACCCAGATCCCAG	5580
Db		5640	AGTGTAGCACTCTTACAGTGCAGGAGCGAGAGCTGGACACTCTCGTAGTACCCAGATCCCAG	5699
QY		5581	AGCCACTAC 5589	
Db		5700	AGCCACTAC 5708	
RESULT 2				
LOCUS	ARL12810	5711 bp	DNA	linear PAT 16-MAY-2001
DEFINITION	Sequence 5 from patent US 6130322.			
ACCESSION	ARL12810			
VERSION	ARL12810.1 GI:14092710			
KEYWORDS				
SOURCE	Unknown.			
ORGANISM	Unclassified.			
REFERENCE	1 (bases 1 to 5711)			
AUTHORS	Murphy,P.D., Allen,A.C.P., Alvares,C.P., Critz,B.S., Olson,S.J., Thunder,D. and Zeng,B. Coding sequences of the human BRCA1 gene Patent: US 6130322-A 5 10-OCT-2000; Location/Qualifiers 1..5711			
JOURNAL				
FEATURES	source	/organism="unknown"		
BASE COUNT	1956 a 1098 c 1274 g 1383 t			
ORIGIN				
Query Match	100.0%; Score 5587.4; DB 6; Length 5711;			
Best Local Similarity	100.0%; Pred. No. 0;			
Matches 5588; Conservative	0; Mismatches 1; Indels 0; Gaps 0;			
QY	1	ATGGA	TTTTATCTGCTCTTCCGGGTTGAGAGACTACAAAATGTCATTTAATGCTATGCACAGAA	60
Db	120	ATGGA	TTTATCTGCTCTTCCGGGTTGAGAGACTACAAAATGTCATTTAATGCTATGCACAGAA	179
QY	61	ATCTTA	GAGTGCOCATCTGTCGTGAGTGTGATCAAGAACGCTGCTCCACAAAGTGTGAC	120
Db	180	ATCTTA	GAGTGCOCATCTGTCGTGAGTGTGATCAAGAACGCTGCTCCACAAAGTGTGAC	239
QY	121	CACATA	TTTTTGCAAAATTTTGGCATGCTGGAACCTTCCAACGAGAGAAAGGCGCTTCACAG	180
Db	240	CACATA	TTTTTGGAATTTTGGCATGCTGGAACCTTCCAACGAGAGAAAGGCGCTTCACAG	299
QY	131	TGTCCT	TTTATGTAGAGATGATATACCAAAGAGAGCTTCAAGAGAAATGACAGATTTTAACT	240
Db	300	TGTCCT	TTTATGTAGAGATGATATACCAAAGAGAGCTTCAAGAGAAATGACAGATTTTAACT	359
QY	241	CACCTGT	TTGGAAGACTATGAAAAATCTTTGTGCTTTTTCAGCTTTGACACAGGTTTGGAG	300
Db	360	CACCTGT	TTGGAAGAGCTATTTGAAAAATCTTTGTGCTTTTTCAGCTTTGACAGAGTTTGGAG	419
QY	301	TATGCA	AACACCTTAATTTTGGCAAAAAAGGAATTAACCTCTCTGGAACATCTTAAAGAT	360
Db	420	TATGCA	AACACCTTAATTTTGGCAAAAAAGGAATTAACCTCTCTGGAACATCTTAAAGAT	479
QY	361	GAACTT	CTATCATTCAGAAAGTATGGGCTTACGAAAACGCTGCAGAAAGACTTCTACAGAGT	420

Db 480 GAAGTTTCATCATCCAAAGTATGGCTACAGAAACCGTGCCAAAGACTTTCACAGAGT 539  
QY 421 GAACCCGAATACCTCTTGCAGGAACAGCTCAGTGTCCACCTCTACCTTGGG 480  
Db 540 GAACCCGAATACCTCTTGCAGGAACAGCTCAGTGTCCACCTCTACCTTGGG 599  
QY 481 ACTGTGAGAACTCTGAGCAAAAGCAGGATACAACTCTAAAAGAGCTGTCTCATTT 540  
Db 600 ACTGTGAGAACTCTGAGCAAAAGCAGGATACAACTCTAAAAGAGCTGTCTCATTT 659  
QY 541 GAATTTGGATCTGATTTCTTCTGAGATACCGTTAATTAAGCAACTTAATTCAGTGGGA 600  
Db 660 GAATTTGGATCTGATTTCTTCTGAGATACCGTTAATTAAGCAACTTAATTCAGTGGGA 719  
QY 601 GATCAAGAAATGTTTACAAATCACCCCTCAGGAACAGGATGAATCACTTTGGAGTCT 660  
Db 720 GATCAAGAAATGTTTACAAATCACCCCTCAGGAACAGGATGAATCACTTTGGAGTCT 779  
QY 661 GCAAAAAAGGCTCTGTGATTTTCTGAGACGATGTAACTAATCTGAACTCATCA 720  
Db 780 GCAAAAAAGGCTCTGTGATTTTCTGAGACGATGTAACTAATCTGAACTCATCA 839  
QY 721 CCCAGTAAATATGATTTGAACACCACTGAGAACGCTGACGTAGAGGATCCAGAAAG 780  
Db 840 CCCAGTAAATATGATTTGAACACCACTGAGAACGCTGACGTAGAGGATCCAGAAAG 899  
QY 781 TATCAGGGTACTCTGTTTCAAACTGCAATGTGAGGATGTGGCAAAATCTCATGCC 840  
Db 900 TATCAGGGTACTCTGTTTCAAACTGCAATGTGAGGATGTGGCAAAATCTCATGCC 959  
QY 841 AGCTCATTCACAGTGAAGACACAGCTTTATTTACTACTAAAGACAAATGAATGTAGAA 900  
Db 960 AGCTCATTCACAGTGAAGACACAGCTTTATTTACTACTAAAGACAAATGAATGTAGAA 1019  
QY 901 AAGGCTGATTTCTGTATATAAAGCAACGCTGCTGTAGCAAGAGCCAACTAAACAGA 960  
Db 1020 AAGGCTGATTTCTGTATATAAAGCAACGCTGCTGTAGCAAGAGCCAACTAAACAGA 1079  
QY 961 TGGGCTGGAAGTAAAGAAACATGATGATGAGGCGACCTCCAGCAAGAAAAAGGTA 1020  
Db 1080 TGGGCTGGAAGTAAAGAAACATGATGATGAGGCGACCTCCAGCAAGAAAAAGGTA 1139  
QY 1021 GATCTGAATGCTGATCCCTGTGTGAGAGAAAAAGATGAATTAAGCAGAACTGCCATGC 1080  
Db 1140 GATCTGAATGCTGATCCCTGTGTGAGAGAAAAAGATGAATTAAGCAGAACTGCCATGC 1199  
QY 1081 TCGAGAAATCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140  
Db 1200 TCGAGAAATCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1259  
QY 1141 AAGGTAATGAGTGGTTTCCAGAAAGTGAACGTGATGATGATGATGATGATGATGATGAT 1200  
Db 1260 AAGGTAATGAGTGGTTTCCAGAAAGTGAACGTGATGATGATGATGATGATGATGATGAT 1319  
QY 1201 GGGGAGTCTGAATCAAAATGCCAAAGTAGTGTGATGATGATGATGATGATGATGATGAT 1260  
Db 1320 GGGGAGTCTGAATCAAAATGCCAAAGTAGTGTGATGATGATGATGATGATGATGATGAT 1379  
QY 1261 GAATATTTCTGTTCTTCAAGAAATAGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
Db 1380 GAATATTTCTGTTCTTCAAGAAATAGATGATGATGATGATGATGATGATGATGATGATGAT 1439  
QY 1321 AATATGTAAGTGAAGAGTGTCTCCAAATCAGTGAAGTGAATTTGAAGCAAAATA 1380  
Db 1440 AATATGTAAGTGAAGAGTGTCTCCAAATCAGTGAAGTGAATTTGAAGCAAAATA 1499  
QY 1381 TTTGGAAAAACCTATCGAAGAAAGCAAGCTCCCACTTAAGCCATTAATGTAATAAT 1440  
Db 1500 TTTGGAAAAACCTATCGAAGAAAGCAAGCTCCCACTTAAGCCATTAATGTAATAAT 1559  
QY 1441 CTAAATATAGAGCAATTTGTACTGAGCCACAGATTAATACAGAGGCTCCCTCACAAAT 1500  
Db 1560 CTAAATATAGAGCAATTTGTACTGAGCCACAGATTAATACAGAGGCTCCCTCACAAAT 1619

QY 1501 AAATTAAGCCTAAAGAGACCTACATCAGGCTTCATCCTGAGATTTTATCAAGAAA 1560  
Db 1620 AAATTAAGCCTAAAGAGACCTACATCAGGCTTCATCCTGAGATTTTATCAAGAAA 1679  
QY 1561 GCAGATTTGGCAGTCTAAAGAGCTCTGAAATGATAATATCAGGGAACCTAACCAACGAG 1620  
Db 1680 GCAGATTTGGCAGTCTAAAGAGCTCTGAAATGATAATATCAGGGAACCTAACCAACGAG 1739  
QY 1621 CAGAAATGTCAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680  
Db 1740 CAGAAATGTCAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1799  
QY 1681 TCTATTCAGATGAGAAAAATCCTAACCCCAATGATGATGATGATGATGATGATGATGAT 1740  
Db 1800 TCTATTCAGATGAGAAAAATCCTAACCCCAATGATGATGATGATGATGATGATGATGAT 1859  
QY 1741 AAACGAAGCTGACCTTAATACAGCAGTATTAAGCAATTAAGCAATTAATTAATATC 1800  
Db 1860 AAACGAAGCTGACCTTAATACAGCAGTATTAAGCAATTAAGCAATTAATTAATATC 1919  
QY 1801 CACAAATTCAAAGCACCCTAAAGAAATAGGCTGAGAGAGAGTCTTCTACAGGCATATT 1860  
Db 1920 CACAAATTCAAAGCACCCTAAAGAAATAGGCTGAGAGAGAGTCTTCTACAGGCATATT 1979  
QY 1861 CATGCGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920  
Db 1980 CATGCGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2039  
QY 1921 ATTGATAGTTGTTCTAGCAGTGAAGATTAAGAAAAAAGTAAACCAACCAATGCCAGTC 1980  
Db 2040 ATTGATAGTTGTTCTAGCAGTGAAGATTAAGAAAAAAGTAAACCAACCAATGCCAGTC 2099  
QY 1981 AGGCACAGCAAAACCTACACATCTATGGAAGTAAAGAACTCTCACTGAGCCAGCAAG 2040  
Db 2100 AGGCACAGCAAAACCTACACATCTATGGAAGTAAAGAACTCTCACTGAGCCAGCAAG 2159  
QY 2041 AGTAACAGCCAAATTAACAGCAAGTAAAGACATGACAGTATCTTCCAGAGCTG 2100  
Db 2160 AGTAACAGCCAAATTAACAGCAAGTAAAGACATGACAGTATCTTCCAGAGCTG 2219  
QY 2101 AAGTAAACAATGACACCTGTTCTTTACTTAAGTGTCAATTAACAGTAAATGAAGAA 2160  
Db 2220 AAGTAAACAATGACACCTGTTCTTTACTTAAGTGTCAATTAACAGTAAATGAAGAA 2279  
QY 2161 TTTGTCAATCTAGCTTCCAGAGAAAGAAAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 2220  
Db 2280 TTTGTCAATCTAGCTTCCAGAGAAAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 2339  
QY 2221 TCTAATTAATGCTGAAGACCCCAAGATCTCATGTTAAGTGAAGAGGTTTGAACAACT 2280  
Db 2340 TCTAATTAATGCTGAAGACCCCAAGATCTCATGTTAAGTGAAGAGGTTTGAACAACT 2399  
QY 2281 GAAAGATCTGTAGAGAGTACAGTATTTTCAATGATGATGATGATGATGATGATGATGATGAT 2340  
Db 2400 GAAAGATCTGTAGAGAGTACAGTATTTTCAATGATGATGATGATGATGATGATGATGATGAT 2459  
QY 2341 GAAAGTATCTGTTACTGGAAGTACACTCTAGGGAAGGCAAAACAGAACCAATTAATA 2400  
Db 2460 GAAAGTATCTGTTACTGGAAGTACACTCTAGGGAAGGCAAAACAGAACCAATTAATA 2519  
QY 2401 TGTGTAGTCAAGTGTGAGCAATTTGAAGCCCAAGGACTAATTTGATGTTTCCAAA 2460  
Db 2520 TGTGTAGTCAAGTGTGAGCAATTTGAAGCCCAAGGACTAATTTGATGTTTCCAAA 2579  
QY 2461 GATATTAAGAAATGACACAGAAAGCTTTAAGTATCATTTGGACATGAAGTTAACACAGT 2520  
Db 2580 GATATTAAGAAATGACACAGAAAGCTTTAAGTATCATTTGGACATGAAGTTAACACAGT 2639  
QY 2521 CGGGAACAAGCTTAAGAAATGGAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2580  
Db 2640 CGGGAACAAGCTTAAGAAATGGAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2699

2581 TTCAGGTTTCAAGAGCGCAGTCATTTGCTGTGTTTCAATCCAGGAAATGCAGAGAG 2640  
|||||  
2700 TTCAGGTTTCAAGAGCGCAGTCATTTGCTGTGTTTCAATCCAGGAAATGCAGAGAG 2759  
2641 GAATGTCGAACATTCCTGCTCCCACTGTGGTCTCTTAAAGAAACAAAGTCCAAAGTCAC 2700  
|||||  
2760 GAATGTCGAACATTCCTGCTCCCACTGTGGTCTCTTAAAGAAACAAAGTCCAAAGTCAC 2819  
2701 TTTGATGTCGAACAAAGAGAGAAATCAAGAGAAATGAGTCTATATCAAGCTCTGA 2760  
|||||  
2820 TTTGATGTCGAACAAAGAGAGAAATCAAGAGAAATGAGTCTATATCAAGCTCTGA 2879  
2761 CAGACAGTTAATATCACTGCTGAGGCTTCTGTGTTGTCAGAAAGATTAAGCAAGTTGAT 2820  
|||||  
2880 CAGACAGTTAATATCACTGCTGAGGCTTCTGTGTTGTCAGAAAGATTAAGCAAGTTGAT 2939  
2821 AATGCCAATGTAGATCAAGAGAGGCTCTAGGTTTGTCTATCATCTCAGTTCAAGGC 2880  
|||||  
2940 AATGCCAATGTAGATCAAGAGAGGCTCTAGGTTTGTCTATCATCTCAGTTCAAGGC 2999  
2881 AACGAACGTGACTCTTACTCCAAATAAACATGAGCTTTTACAAACCATATCGTATA 2940  
|||||  
3000 AACGAACGTGACTCTTACTCCAAATAAACATGAGCTTTTACAAACCATATCGTATA 3059  
2941 CCACCACTTTTCCCATCAAGTCATTTGTTAAACTTAATGTAGAAATTCGTCTAGAG 3000  
|||||  
3060 CCACCACTTTTCCCATCAAGTCATTTGTTAAACTTAATGTAGAAATTCGTCTAGAG 3119  
3001 GAAACCTTTGAGGAACTCATGTGACCTGAAAGAAATGGAAATGGAACATTTCCA 3060  
|||||  
3120 GAAACCTTTGAGGAACTCATGTGACCTGAAAGAAATGGAAATGGAACATTTCCA 3179  
3061 AGTAGAGTGAAGCAATTAAGCCGTAAATACATTAGAGAAATGTTTAAAGAGCCAGC 3120  
|||||  
3180 AGTAGAGTGAAGCAATTAAGCCGTAAATACATTAGAGAAATGTTTAAAGAGCCAGC 3239  
3121 TCACCAATATTAAAGAGTGAAGTTCAGTACTAATGAAGTGGCTCCAGTATTAAATGA 3180  
|||||  
3240 TCACCAATATTAAAGAGTGAAGTTCAGTACTAATGAAGTGGCTCCAGTATTAAATGA 3299  
3181 ATAGTATTCAGTGAAGAAACATTCAGAGAGAACTGATGAAGAGAGGCGCAAAATG 3240  
|||||  
3300 ATAGTATTCAGTGAAGAAACATTCAGAGAGAACTGATGAAGAGAGGCGCAAAATG 3359  
3241 AATGCTATGCTAGATTAGGGGTTTGGCAACTGAGGCTATTAACAAAGTCTTCCGTGA 3300  
|||||  
3360 AATGCTATGCTAGATTAGGGGTTTGGCAACTGAGGCTATTAACAAAGTCTTCCGTGA 3419  
3301 AGTAATTTGAGCATCTGGAATTAAGAAAGCAAGAAATATGAAGAGTACTGACTGTT 3360  
|||||  
3420 AGTAATTTGAGCATCTGGAATTAAGAAAGCAAGAAATATGAAGAGTACTGACTGTT 3479  
3361 AATACAGATTTCTCCATATGATGATTTAGATTAAGTGAACAGCTATGGGAATGAT 3420  
|||||  
3480 AATACAGATTTCTCCATATGATGATTTAGATTAAGTGAACAGCTATGGGAATGAT 3539  
3421 CATGATCTCAGGTTTGTTCAGACACCTGATGACCTGTAGATGATGATGATGAATGA 3480  
|||||  
3540 CATGATCTCAGGTTTGTTCAGACACCTGATGACCTGTAGATGATGATGATGAATGA 3599  
3481 GAAGTACTAGTTTGTCTGGAATGACATTAAGGAAGTTGCTGTTTTAGCAAAAGC 3540  
|||||  
3600 GAAGTACTAGTTTGTCTGGAATGACATTAAGGAAGTTGCTGTTTTAGCAAAAGC 3659  
3541 GTCCAGAAAGAGAGAGAGAGGAGTCTAGGCTTACACCATACATTTGGCTGAG 3600  
|||||  
3660 GTCCAGAAAGAGAGAGAGGAGTCTAGGCTTACACCATACATTTGGCTGAG 3719  
3601 GGTTCAGAGAGAGGAGAGAGGAGTCTAGGCTTACAGAGAGAGTATATGATGAGAT 3660  
|||||  
3720 GGTTCAGAGAGAGGAGAGAGGAGTCTAGGCTTACAGAGAGAGTATATGATGAGAT 3779  
3661 GAAGAGCTTCCCTGCTCCAAACATTTGTTTGTAAAGTAACATATACCTTCTGAG 3720

|||||  
3780 GAAGAGCTTCCCTGCTCCAAACATTTGTTTGTAAAGTAACAAATATACCTTCTGAG 3839  
3721 TCTACTAGGATATGACACCGTTGCTACCGAGTGTCTGTCTTAAAGACAGAGAGAAATTA 3780  
|||||  
3840 TCTACTAGGATATGACACCGTTGCTACCGAGTGTCTGTCTTAAAGACAGAGAGAAATTA 3899  
3781 TTATCATTTGAGAAATGATTAATGATGATGATGATGATGATGATGATGATGATGAT 3840  
|||||  
3900 TTATCATTTGAGAAATGATTAATGATGATGATGATGATGATGATGATGATGATGAT 3959  
3841 CAGGAACATCACCCTTATGTAAGAGAAACAAATGTTCTGCTAGCTTTTCTTACAGTGC 3900  
|||||  
3960 CAGGAACATCACCCTTATGTAAGAGAAACAAATGTTCTGCTAGCTTTTCTTACAGTGC 4019  
3901 AGTGAATTTGGAAGACTGTGACTGCAATTAACAAACCCAGAGATCTTCTGTTGTTGTTCT 3960  
|||||  
4020 AGTGAATTTGGAAGACTGTGACTGCAATTAACAAACCCAGAGATCTTCTGTTGTTGTTCT 4079  
3961 TCCAAACAAATGAGGATCAGTCTGAAACCCAGGAGTTGCTGTGAGTGAAGAGAAATG 4020  
|||||  
4080 TCCAAACAAATGAGGATCAGTCTGAAACCCAGGAGTTGCTGTGAGTGAAGAGAAATG 4139  
4021 GTTTCAGATGATGAGAAAGAGAGAGGCTTGGAGAAATATATCAAGAGAGCAAGC 4080  
|||||  
4140 GTTTCAGATGATGAGAAAGAGAGAGGCTTGGAGAAATATATCAAGAGAGCAAGC 4199  
4081 ATGGAATTTCAACTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4140  
|||||  
4200 ATGGAATTTCAACTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4259  
4141 GACTGCTCAGGAGCTATCTCTCAGAGTGCATTTTAAACACTCAGCAGAGAGGATACATG 4200  
|||||  
4260 GACTGCTCAGGAGCTATCTCTCAGAGTGCATTTTAAACACTCAGCAGAGAGGATACATG 4319  
4201 CAACATTAACCTGATTAAGCTCCAGCAGAGAAATGCTGAATGAAAGTGTGTAGAACG 4260  
|||||  
4320 CAACATTAACCTGATTAAGCTCCAGCAGAGAAATGCTGAATGAAAGTGTGTAGAACG 4379  
4261 CATGGAGGAGGAGCTTCTTAACAGCTACCTTCAATCATTAAGTACTTCTGCTGAG 4320  
|||||  
4380 CATGGAGGAGGAGCTTCTTAACAGCTACCTTCAATCATTAAGTACTTCTGCTGAG 4439  
4321 GACCTGCGAAATCCAGAACAGACATCAGAAAGAGAGTATTAACCTTCAAGAAAGT 4380  
|||||  
4440 GACCTGCGAAATCCAGAACAGACATCAGAAAGAGAGTATTAACCTTCAAGAAAGT 4499  
4381 AGTGAATTTCACTTAAGCCAGAAATCCAGAGAGGCTTCTGCTGACAAATTTGAGTGTCT 4440  
|||||  
4500 AGTGAATTTCACTTAAGCCAGAAATCCAGAGAGGCTTCTGCTGACAAATTTGAGTGTCT 4559  
4441 GCAGATAGTTCTAACAGTAAATTAAGAACAGAGAGTGAAGGTCATCCCTTCTAAA 4500  
|||||  
4560 GCAGATAGTTCTAACAGTAAATTAAGAACAGAGAGTGAAGGTCATCCCTTCTAAA 4619  
4501 TGCCCATCATATGATGATATGATGATGATGATGATGATGATGATGATGATGATGAT 4560  
|||||  
4620 TGCCCATCATATGATGATATGATGATGATGATGATGATGATGATGATGATGATGAT 4679  
4561 AACTACCCATCTCAAGAGAGAGTCAATTAAGTGTGTGATGTGAGAGAGCAACACTGGA 4620  
|||||  
4680 AACTACCCATCTCAAGAGAGAGTCAATTAAGTGTGTGATGTGAGAGAGCAACACTGGA 4739  
4621 GAGTCTGGGCGACAGAGATTTGCGGAACATTTCTTCCCAAGGCAAGATCTGAGGGA 4680  
|||||  
4740 GAGTCTGGGCGACAGAGATTTGCGGAACATTTCTTCCCAAGGCAAGATCTGAGGGA 4799  
4681 ACCCTTACCTGGAATCTGGAATCAGCTTCTCTCTGATGAGACCTGATGATCTTCT 4740  
|||||  
4800 ACCCTTACCTGGAATCTGGAATCAGCTTCTCTCTGATGAGACCTGATGATCTTCT 4859  
4741 GAAGACAGAGCCCAAGAGTCACTGCTGTGGCAACATACATCTTCAACCTTGCATG 4800  
|||||



Db	4860	GAAGACAGAGAGCCCAAGAGTACGCTCGTGTTGGACAAATACCATCTTCAACCTCGCATGTG	49119
QY	4801	AAAGTTCCTCCCAATTGAAAGTTGCCAGAAATCTGCCAGAGTCAGCTGCTGCTCACTACT	48606
Db	4920	AAAGTTCCTCCCAATTGAAAGTTGCCAGAAATCTGCCAGAGTCAGCTGCTGCTCACTACT	49799
QY	4861	GATACCTGCTGGGTATTAATGCAATGGAGAAAGTGTGAGCAGAGGAGAAAGCCAGAAATTGACA	49220
Db	4980	GATACCTGCTGGGTATTAATGCAATGGAGAAAGTGTGAGCAGAGGAGAAAGCCAGAAATTGACA	50399
QY	4921	GCTTCAACAGAAAGGGTCCACAAAAGAAATGTCATGTGGTGGTGTGGCTCACCCACAGA	49808
Db	5040	GCTTCAACAGAAAGGGTCCACAAAAGAAATGTCATGTGGTGGTGTGGCTCACCCACAGA	50999
QY	4981	GAATTTATGCTCGTGTACAAAGTTTGCAGAAACACCAATCACTTTAACTAAATCTAAT	50404
Db	5100	GAATTTATGCTCGTGTACAAAGTTTGCAGAAACACCAATCACTTTAACTAAATCTAAT	51599
QY	5041	ACTAGAGAGCACTACTCATGTTGTTATGAAAAACAGATGCTGAGTTTGTGTGAAACGAGACA	51000
Db	5160	ACTAGAGAGCACTACTCATGTTGTTATGAAAAACAGATGCTGAGTTTGTGTGAAACGAGACA	52199
QY	5101	CTGAATATTTTTCATGAGAAATGGCGGAGGAAATGGGAGTACTACTTTCCTGGGTGACC	51606
Db	5220	CTGAATATTTTTCATGAGAAATGGCGGAGGAAATGGGAGTACTACTTTCCTGGGTGACC	52799
QY	5161	CAGTCTATTTAAAGAAAGAAAAAAGTCTGAGTACGATGTTTGAAGTCAGAGAGAGATGTG	52204
Db	5280	CAGTCTATTTAAAGAAAGAAAAAAGTCTGAGTACGATGTTTGAAGTCAGAGAGAGATGTG	53399
QY	5221	GTCATATGAGAAACACCAAGGTCCAAAGCGAGCAGAGATGCCAGAGACAGAAAGATC	52806
Db	5340	GTCATATGAGAAACACCAAGGTCCAAAGCGAGCAGAGATGCCAGAGACAGAAAGATC	53999
QY	5281	TTCAAGGGGGCTAGAAATCTGTGCTATGAGGGCCCTTACCAACATGGCCACAGATCAACTG	53404
Db	5400	TTCAAGGGGGCTAGAAATCTGTGCTATGAGGGCCCTTACCAACATGGCCACAGATCAACTG	54599
QY	5341	GAATGGATGTGACAGCTGTGTGTCTCTGTGTGTAAGAGAGCTTTGATCATCTACCCCTT	54004
Db	5460	GAATGGATGTGACAGCTGTGTGTCTCTGTGTGTAAGAGAGCTTTGATCATCTACCCCTT	55199
QY	5401	GGCACAGGTGTCCACCCCAATTTGTGGTTGTGACGCACAGATGCTTGAGACAGAGCAATGGC	54606
Db	5520	GGCACAGGTGTCCACCCCAATTTGTGGTTGTGACGCACAGATGCTTGAGACAGAGCAATGGC	55799
QY	5461	TTTCATGCAATTTGGGACAGATGTGTGAGGCACCTGTGTGTGAGCCCAAGATGGGTGTGAC	55204
Db	5580	TTTCATGCAATTTGGGACAGATGTGTGAGGCACCTGTGTGTGAGCCCAAGATGGGTGTGAC	56399
QY	5521	AGTTAGCACTCTACAGATGTCAGAGAGTGGACACCTACTAGTATACCCACAGATCCCCAC	55806
Db	5640	AGTTAGCACTCTACAGATGTCAGAGAGTGGACACCTACTAGTATACCCACAGATCCCCAC	56999
QY	5581	AGCCACTAC 5589	
Db	5700	AGCCACTAC 5708	
RESULT 3			
LOCUS	AR033056	5711 bp	DNA
DEFINITION	Sequence 1 from patent US 5869245.	linear	PAT 29-SEP-1999
ACCESSION	AR033056		
VERSION	AR033056.1	GI:5948661	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 5711)		
AUTHORS	Yeung,A.T.		
TITLE	Mismatch endonuclease and its use in identifying mutations in targeted polynucleotide strands		

JOURNAL		Patent: US 5869245-A 1 09-FEB-1999:	
FEATURES		Location/Qualifiers	
source	1..5711		
BASE COUNT	1956 a 1099 c 1274 g 1382 t		
ORIGIN			
Query Match	99.9%;	Score 5585.8;	DB 6; Length 5711;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 5587;	Conservative 0;	Mismatches 2;	Indels 0; Gaps 0
QY	1 ATGATTTTATCTGCTCTTCGCGCTTAAAGAGTACAAATATGCTATTATGCTATGCGAAA	60	
DB	120 ATGATTTTATCTGCTCTTCGCGCTTAAAGAGTACAAATATGCTATTATGCTATGCGAAA	179	
QY	61 ATCTTAGAGTGTCCCATCTGCTGTGGAGTTGATCAAGGAACCTGTCTCCCAAGAGTGTGAC	120	
DB	180 ATCTTAGAGTGTCCCATCTGCTGTGGAGTTGATCAAGGAACCTGTCTCCCAAGAGTGTGAC	239	
QY	121 CACATATTTTGCAGATTTTGGCATGCTGAGAACTTCTCAACAGAGAAAGGGCTTCACAG	180	
DB	240 CACATATTTTGCAGATTTTGGCATGCTGAGAACTTCTCAACAGAGAAAGGGCTTCACAG	299	
QY	181 TGTCTTTATGTAGAAATGATATTAACCAAGAGGCTTCAAGAAAGTACGAGATTAGT	240	
DB	300 TGTCTTTATGTAGAAATGATATTAACCAAGAGGCTTCAAGAAAGTACGAGATTAGT	359	
QY	241 CACCTGTTGGAAGACCTATGTAAGAAATCAATTTGCTTTGACCTTGACAGACAGTTTGGAG	300	
DB	360 CACCTGTTGGAAGACCTATGTAAGAAATCAATTTGCTTTGACCTTGACAGACAGTTTGGAG	419	
QY	301 TATGCAACAGCTATATATTTTGGCAAAAAAGSAAATTAACCTCTCTGAAACATCTTAAAGAT	360	
DB	420 TATGCAACAGCTATATATTTTGGCAAAAAAGSAAATTAACCTCTCTGAAACATCTTAAAGAT	479	
QY	361 GAACTTTTATCATATCCAAAGTATGGGCTTACAGAAACCGTGGCAAAAGACTTTACAGAGT	420	
DB	480 GAACTTTTATCATATCCAAAGTATGGGCTTACAGAAACCGTGGCAAAAGACTTTACAGAGT	539	
QY	421 GAACCCGAAATTCCTTCCTGGAGAAACAGTCTCAGTGTCAACTCTCTTAACCTTGA	480	
DB	540 GAACCCGAAATTCCTTCCTGGAGAAACAGTCTCAGTGTCAACTCTCTTAACCTTGA	599	
QY	481 ACTGTGAGACCTGTGAGGACCAAGAGGGGATACACACCTCAAAAGACGCTGTCTTACAT	540	
DB	600 ACTGTGAGACCTGTGAGGACCAAGAGGGGATACACACCTCAAAAGACGCTGTCTTACAT	659	
QY	541 GAATTGGGATCTGATTTCTTCTGAGAGTACCGTTAATTAAGCAACTTATTTGACGTGGGA	600	
DB	660 GAATTGGGATCTGATTTCTTCTGAGAGTACCGTTAATTAAGCAACTTATTTGACGTGGGA	719	
QY	601 GATCAAGAAATGTTCACAAATCACCCCTCAAGGAACCAAGGATGAANAATCAGTTGGATTCT	660	
DB	720 GATCAAGAAATGTTCACAAATCACCCCTCAAGGAACCAAGGATGAANAATCAGTTGGATTCT	779	
QY	661 GCAAAAAAGCGCTGTGAAATTTCTGAGAGGAGTGTACAAATACTAATCATCATCA	720	
DB	780 GCAAAAAAGCGCTGTGAAATTTCTGAGAGGAGTGTACAAATACTAATCATCATCA	839	
QY	721 CCCAGTAATTAATGATTTTGAAACCACTGAGAAAGCGTGCAGTACAGGCGATCCAGAAAG	780	
DB	840 CCCAGTAATTAATGATTTTGAAACCACTGAGAAAGCGTGCAGTACAGGCGATCCAGAAAG	899	
QY	781 TATCAGGTAATGTCTGTTTCAAACTTGGATGTGAGAGCCATGTGGCAAAATACTCATGCC	840	
DB	900 TATCAGGTAATGTCTGTTTCAAACTTGGATGTGAGAGCCATGTGGCAAAATACTCATGCC	959	
QY	841 AGCTCATTTACAGCATGAGACAGCGCTTTATATCTCACATTAAGAGACAGATGATGTAGAA	900	
DB	960 AGCTCATTTACAGCATGAGACAGCGCTTTATATCTCACATTAAGAGACAGATGATGTAGAA	1011	
QY	901 AAGCGTAATTTCTGTAATTAAGCAACAGCGCTGCTTAGCAAGGAGGCCAATCAACGA	960	

Dh 1020 AAGGCTGAATCTGTAATAAAAGCAAAACGCTGGCTTAGCAAGAGCCAACTAATACAG 1079  
Qy 961 TGGGCTGGAAGTAAGAAACATGTAAATGATGGCGGACATCCAGCAGCAAGAAAAAGTA 1020  
Dh 1080 TGGGCTGGAGTAAGAAACATGTAAATGATGGCGGACATCCAGCAGCAAGAAAAAGTA 1139  
Qy 1021 GATCGAATGCTGATCCCTGTGTGAGAGAAAAAATGGAATAGCAGAAACTGCCATC 1080  
Dh 1140 GATCGAATGCTGATCCCTGTGTGAGAGAAAAAATGGAATAGCAGAAACTGCCATC 1199  
Qy 1081 TCAGAGAACTCTAGAGATCTGAGAAATGTTCTTGATACACTAAATAGCAGATTCA 1140  
Dh 1200 TCAGAGAACTCTAGAGATCTGAGAAATGTTCTTGATACACTAAATAGCAGATTCA 1259  
Qy 1141 AAAGTAATGATGGGTTTCCAGAGTGAACCTGTAGGTTCTGATGCTCAGATGAT 1200  
Dh 1280 AAAGTAATGATGGGTTTCCAGAGTGAACCTGTAGGTTCTGATGCTCAGATGAT 1319  
Qy 1201 GGGGAGTCTGAATCAAAATGCCAAAGTAGCTGATGATTTGGAGCTTCTAAATAGTAGAT 1260  
Dh 1320 GGGGAGTCTGAATCAAAATGCCAAAGTAGCTGATGATTTGGAGCTTCTAAATAGTAGAT 1379  
Qy 1261 GAATATTTCTGCTTCTCAGAGAAATAGACTTACTGCGCAGTATCTCTATGAGCTTTA 1320  
Dh 1380 GAATATTTCTGCTTCTCAGAGAAATAGACTTACTGCGCAGTATCTCTATGAGCTTTA 1439  
Qy 1321 ATATGTAAAGTGAAGAGTTCACCTCCAAATCAGTAGAGATTAATTTGAAGCAAAATA 1380  
Dh 1440 ATATGTAAAGTGAAGAGTTCACCTCCAAATCAGTAGAGATTAATTTGAAGCAAAATA 1499  
Qy 1391 TTTGGGAAAACCTATCGAAGAGAGCCCTCCCACTTAAGCCATTAAGCTTAAGCTTA 1440  
Dh 1500 TTTGGGAAAACCTATCGAAGAGAGCCCTCCCACTTAAGCCATTAAGCTTAAGCTTA 1559  
Qy 1441 CTAATTTAAGAGCATTTTGTCTAGAGCCAGATATATCAAGAGGCTCCCTCAGAAAT 1500  
Dh 1560 CTAATTTAAGAGCATTTTGTCTAGAGCCAGATATATCAAGAGGCTCCCTCAGAAAT 1619  
Qy 1501 AAATTAAGAGCATTTAAGAGAGACCTATCAGAGCTTCACTGAGAGTTTATCAAGAA 1560  
Dh 1620 AAATTAAGAGCATTTAAGAGAGACCTATCAGAGCTTCACTGAGAGTTTATCAAGAA 1679  
Qy 1561 GCAGATTTGGCAGTTCAAAAGAGCTCTCGAATATGATTAATCAGGAACTTAACCAACGAG 1620  
Dh 1680 GCAGATTTGGCAGTTCAAAAGAGCTCTCGAATATGATTAATCAGGAACTTAACCAACGAG 1739  
Qy 1621 CAGATGTCAGATGATGAATTTACTATAGTGTGTCATGAGATTAACCAAAAGGTGAT 1680  
Dh 1740 CAGATGTCAGATGATGAATTTACTATAGTGTGTCATGAGATTAACCAAAAGGTGAT 1799  
Qy 1681 TCTATTCAGATGAGAAAAATCTTAACCAATAGAACTCGAAAAAAGATCTGCTTC 1740  
Dh 1800 TCTATTCAGATGAGAAAAATCTTAACCAATAGAACTCGAAAAAAGATCTGCTTC 1859  
Qy 1741 AAAACGAAGCTGAACCTATTAAGAGAGCATATAGCAATATGGAATCGAATTAATATC 1800  
Dh 1860 AAAACGAAGCTGAACCTATTAAGAGAGCATATAGCAATATGGAATCGAATTAATATC 1919  
Qy 1801 CACATTTAAAAAGCACTTAAGAAATAGAGTGTGAGAGAGAACTCTTCTACCGAGCATAT 1860  
Dh 1920 CACATTTAAAAAGCACTTAAGAAATAGAGTGTGAGAGAGAACTCTTCTACCGAGCATAT 1979  
Qy 1861 CATGGCTTTGAACCTAGTAGTCTAGTAATCTAAGCCACCTAATTTGTCTGAAATGCA 1920  
Dh 1980 CATGGCTTTGAACCTAGTAGTCTAGTAATCTAAGCCACCTAATTTGTCTGAAATGCA 2039  
Qy 1921 ATTGATGTTGTTCTAGAGAGTGAAGATTAAGAAAAAAGTACAACTTAAGCCAGTC 1980  
Dh 2040 ATTGATGTTGTTCTAGAGAGTGAAGATTAAGAAAAAAGTACAACTTAAGCCAGTC 2099  
Qy 1981 AGGCACACAGAAACCTCACTCACTGAGAGTAAAGAACTGCAACTGAGCCAGAAAG 2040  
Dh 2100 AGGCACACAGAAACCTCACTCACTGAGAGTAAAGAACTGCAACTGAGCCAGAAAG 2159

Qy 2041 AGTAACAGCCAAATGAAAGAGACAGATGAAAGACATGACAGTACTTTCCAGAGCT 2100  
Dh 2160 AGTAACAGCCAAATGAAAGAGACAGATGAAAGACATGACAGTACTTTCCAGAGCT 2219  
Qy 2101 AAGTTAACAATGCAACCTGGTCTTTACTAGTGTCAATATACAGTGAACCTTAAGAA 2160  
Dh 2220 AAGTTAACAATGCAACCTGGTCTTTACTAGTGTCAATATACAGTGAACCTTAAGAA 2279  
Qy 2161 TTTGTCAATCTTACCTTCCAGAGAGAAAAAGAAAGAACTTGAACAGTTAAAGTG 2220  
Dh 2280 TTTGTCAATCTTACCTTCCAGAGAGAAAAAGAAAGAACTTGAACAGTTAAAGTG 2339  
Qy 2221 TCTAATATGCTGAAGAGCCCAAGATGCTAATGAGAGAAAGGCTTTGCAAACT 2280  
Dh 2340 TCTAATATGCTGAAGAGCCCAAGATGCTAATGAGAGAAAGGCTTTGCAAACT 2399  
Qy 2281 GAAAGATCTGTAGAGAGTGAAGATTTTCAATTTGTAAGTGTATGATGAGCACTCAG 2340  
Dh 2400 GAAAGATCTGTAGAGAGTGAAGATTTTCAATTTGTAAGTGTATGATGAGCACTCAG 2459  
Qy 2341 GAAAGATCTGTAGAGAGTGAAGATTTTCAATTTGTAAGTGTATGATGAGCACTCAG 2400  
Dh 2460 GAAAGATCTGTAGAGAGTGAAGATTTTCAATTTGTAAGTGTATGATGAGCACTCAG 2519  
Qy 2401 TGTGTAGTCACTGTGACAGCATTTGAAAAACCCAGAGGACTAATTCATGTTGTTCCAA 2460  
Dh 2520 TGTGTAGTCACTGTGACAGCATTTGAAAAACCCAGAGGACTAATTCATGTTGTTCCAA 2579  
Qy 2461 GATTAATGAAATGACACAGAGGCTTTAAGTATTCATTTGGAGATGAAGTTAACCAGAT 2520  
Dh 2580 GATTAATGAAATGACACAGAGGCTTTAAGTATTCATTTGGAGATGAAGTTAACCAGAT 2639  
Qy 2521 CGGGAACAGAGATGAAGATGAAGAAAGTGAACCTGATGCTCAGTATTTGCCAATTA 2580  
Dh 2640 CGGGAACAGAGATGAAGATGAAGAAAGTGAACCTGATGCTCAGTATTTGCCAATTA 2639  
Qy 2581 TTTCAAGGTTTCAAAAGGCGCAGTCACTGCTGTTTCAATTCAGAAATGCGAAGAG 2640  
Dh 2700 TTTCAAGGTTTCAAAAGGCGCAGTCACTGCTGTTTCAATTCAGAAATGCGAAGAG 2759  
Qy 2641 GAATGTCAGATCTTCTGCCACTCTGCGGCTTTAAAGAAACAAAGTCCAAAGTCACT 2700  
Dh 2760 GAATGTCAGATCTTCTGCCACTCTGCGGCTTTAAAGAAACAAAGTCCAAAGTCACT 2819  
Qy 2701 TTTGAATGGAACAAAGAGAAAGAAATCAAGAAAGATGATCAATATCAACCTGTA 2760  
Dh 2820 TTTGAATGGAACAAAGAGAAAGAAATCAAGAAAGATGATCAATATCAACCTGTA 2879  
Qy 2761 CAGACAGTTAATATCACTGCAAGGCTTCTGCTGTTGTCAGAAAGATTAAGCCAGTAT 2820  
Dh 2880 CAGACAGTTAATATCACTGCAAGGCTTCTGCTGTTGTCAGAAAGATTAAGCCAGTAT 2939  
Qy 2821 AATGCCAAATGATATCAAGAGAGGCTCTAGAGTTTGTCTATCATCTCAGTCAAGGC 2880  
Dh 2940 AATGCCAAATGATATCAAGAGAGGCTCTAGAGTTTGTCTATCATCTCAGTCAAGGC 2939  
Qy 2881 AAGCAATGAGAGCTATCACTCAAAATTAAGATGAGCTTTACAAACCCATATCGATA 2940  
Dh 3000 AAGCAATGAGAGCTATCACTCAAAATTAAGATGAGCTTTACAAACCCATATCGATA 3059  
Qy 2941 CCAGCAGTTTCCCATCAAGTCAATTTGTTAAACCTAAGAAATTCCTGAGAG 3000  
Dh 3060 CCAGCAGTTTCCCATCAAGTCAATTTGTTAAACCTAAGAAATTCCTGAGAG 3119  
Qy 3001 GAAAACTTTGAGGAACATTCATGACCTGAAAGAGAAATGGAATGGAACATTCGA 3060  
Dh 3120 GAAAACTTTGAGGAACATTCATGACCTGAAAGAGAAATGGAATGGAACATTCGA 3179  
Qy 3061 AGTACAGTGAACAAATTAAGCCGTAATTAAGTGAAGAAATGTTTAAAGAAAGCCAGC 3120  
Dh 3180 AGTACAGTGAACAAATTAAGCCGTAATTAAGTGAAGAAATGTTTAAAGAAAGCCAGC 3239

3121 TCACCAATATTAAAGAGTAGCTTCAGTACTAATGAAGTGGCTCCAGTATTAATGA 3180  
|||||  
3240 TCACCAATATTAAAGAGTAGCTTCAGTACTAATGAAGTGGCTCCAGTATTAATGA 3299  
|||||  
3181 ATAGGTTCAGTGAAGAAACATTCAGCAGAACTAGGTAGAAACAGAGGGCCAAATTTG 3240  
|||||  
3300 ATAGGTTCAGTGAAGAAACATTCAGCAGAACTAGGTAGAAACAGAGGGCCAAATTTG 3359  
|||||  
3241 ATAGGTTCAGTGAAGAAACATTCAGCAGAACTAGGTAGAAACAGAGGGCCAAATTTG 3300  
|||||  
3360 ATAGGTTCAGTGAAGAAACATTCAGCAGAACTAGGTAGAAACAGAGGGCCAAATTTG 3419  
|||||  
3301 ATAGGTTCAGTGAAGAAACATTCAGCAGAACTAGGTAGAAACAGAGGGCCAAATTTG 3360  
|||||  
3420 ATAGGTTCAGTGAAGAAACATTCAGCAGAACTAGGTAGAAACAGAGGGCCAAATTTG 3479  
|||||  
3361 ATAGGTTCAGTGAAGAAACATTCAGCAGAACTAGGTAGAAACAGAGGGCCAAATTTG 3420  
|||||  
3480 ATAGGTTCAGTGAAGAAACATTCAGCAGAACTAGGTAGAAACAGAGGGCCAAATTTG 3539  
|||||  
3421 ATAGGTTCAGTGAAGAAACATTCAGCAGAACTAGGTAGAAACAGAGGGCCAAATTTG 3480  
|||||  
3540 ATAGGTTCAGTGAAGAAACATTCAGCAGAACTAGGTAGAAACAGAGGGCCAAATTTG 3599  
|||||  
3481 ATAGGTTCAGTGAAGAAACATTCAGCAGAACTAGGTAGAAACAGAGGGCCAAATTTG 3540  
|||||  
3600 ATAGGTTCAGTGAAGAAACATTCAGCAGAACTAGGTAGAAACAGAGGGCCAAATTTG 3659  
|||||  
3541 ATAGGTTCAGTGAAGAAACATTCAGCAGAACTAGGTAGAAACAGAGGGCCAAATTTG 3600  
|||||  
3660 ATAGGTTCAGTGAAGAAACATTCAGCAGAACTAGGTAGAAACAGAGGGCCAAATTTG 3719  
|||||  
3601 ATAGGTTCAGTGAAGAAACATTCAGCAGAACTAGGTAGAAACAGAGGGCCAAATTTG 3660  
|||||  
3720 ATAGGTTCAGTGAAGAAACATTCAGCAGAACTAGGTAGAAACAGAGGGCCAAATTTG 3779  
|||||  
3661 ATAGGTTCAGTGAAGAAACATTCAGCAGAACTAGGTAGAAACAGAGGGCCAAATTTG 3720  
|||||  
3780 ATAGGTTCAGTGAAGAAACATTCAGCAGAACTAGGTAGAAACAGAGGGCCAAATTTG 3839  
|||||  
3721 ATAGGTTCAGTGAAGAAACATTCAGCAGAACTAGGTAGAAACAGAGGGCCAAATTTG 3780  
|||||  
3840 ATAGGTTCAGTGAAGAAACATTCAGCAGAACTAGGTAGAAACAGAGGGCCAAATTTG 3899  
|||||  
3781 ATAGGTTCAGTGAAGAAACATTCAGCAGAACTAGGTAGAAACAGAGGGCCAAATTTG 3840  
|||||  
3900 ATAGGTTCAGTGAAGAAACATTCAGCAGAACTAGGTAGAAACAGAGGGCCAAATTTG 3959  
|||||  
3841 ATAGGTTCAGTGAAGAAACATTCAGCAGAACTAGGTAGAAACAGAGGGCCAAATTTG 3900  
|||||  
3960 ATAGGTTCAGTGAAGAAACATTCAGCAGAACTAGGTAGAAACAGAGGGCCAAATTTG 4019  
|||||  
3901 ATAGGTTCAGTGAAGAAACATTCAGCAGAACTAGGTAGAAACAGAGGGCCAAATTTG 3960  
|||||  
4020 ATAGGTTCAGTGAAGAAACATTCAGCAGAACTAGGTAGAAACAGAGGGCCAAATTTG 4079  
|||||  
3961 ATAGGTTCAGTGAAGAAACATTCAGCAGAACTAGGTAGAAACAGAGGGCCAAATTTG 4020  
|||||  
4080 ATAGGTTCAGTGAAGAAACATTCAGCAGAACTAGGTAGAAACAGAGGGCCAAATTTG 4139  
|||||  
4021 ATAGGTTCAGTGAAGAAACATTCAGCAGAACTAGGTAGAAACAGAGGGCCAAATTTG 4080  
|||||  
4140 ATAGGTTCAGTGAAGAAACATTCAGCAGAACTAGGTAGAAACAGAGGGCCAAATTTG 4199  
|||||  
4081 ATAGGTTCAGTGAAGAAACATTCAGCAGAACTAGGTAGAAACAGAGGGCCAAATTTG 4140  
|||||  
4200 ATAGGTTCAGTGAAGAAACATTCAGCAGAACTAGGTAGAAACAGAGGGCCAAATTTG 4259  
|||||  
4141 ATAGGTTCAGTGAAGAAACATTCAGCAGAACTAGGTAGAAACAGAGGGCCAAATTTG 4200  
|||||  
4260 ATAGGTTCAGTGAAGAAACATTCAGCAGAACTAGGTAGAAACAGAGGGCCAAATTTG 4319  
|||||  
4201 ATAGGTTCAGTGAAGAAACATTCAGCAGAACTAGGTAGAAACAGAGGGCCAAATTTG 4260  
|||||

|||||  
4320 CAACATTAACCTGATAAGAGTCCAGCAGAAATGGCTGAGTACAGAGTGTAGAACAG 4379  
|||||  
4261 CATGGAGCAGACCTCTTCAACAGCTACACCTCTTCAATAGTACAGTCTTCTGAG 4320  
|||||  
4380 CATGGAGCAGACCTCTTCAACAGCTACACCTCTTCAATAGTACAGTCTTCTGAG 4439  
|||||  
4321 CATGGAGCAGACCTCTTCAACAGCTACACCTCTTCAATAGTACAGTCTTCTGAG 4380  
|||||  
4440 CATGGAGCAGACCTCTTCAACAGCTACACCTCTTCAATAGTACAGTCTTCTGAG 4499  
|||||  
4381 CATGGAGCAGACCTCTTCAACAGCTACACCTCTTCAATAGTACAGTCTTCTGAG 4440  
|||||  
4500 CATGGAGCAGACCTCTTCAACAGCTACACCTCTTCAATAGTACAGTCTTCTGAG 4559  
|||||  
4441 CATGGAGCAGACCTCTTCAACAGCTACACCTCTTCAATAGTACAGTCTTCTGAG 4500  
|||||  
4560 CATGGAGCAGACCTCTTCAACAGCTACACCTCTTCAATAGTACAGTCTTCTGAG 4619  
|||||  
4501 CATGGAGCAGACCTCTTCAACAGCTACACCTCTTCAATAGTACAGTCTTCTGAG 4560  
|||||  
4620 CATGGAGCAGACCTCTTCAACAGCTACACCTCTTCAATAGTACAGTCTTCTGAG 4679  
|||||  
4561 CATGGAGCAGACCTCTTCAACAGCTACACCTCTTCAATAGTACAGTCTTCTGAG 4620  
|||||  
4680 CATGGAGCAGACCTCTTCAACAGCTACACCTCTTCAATAGTACAGTCTTCTGAG 4739  
|||||  
4621 CATGGAGCAGACCTCTTCAACAGCTACACCTCTTCAATAGTACAGTCTTCTGAG 4680  
|||||  
4740 CATGGAGCAGACCTCTTCAACAGCTACACCTCTTCAATAGTACAGTCTTCTGAG 4799  
|||||  
4681 CATGGAGCAGACCTCTTCAACAGCTACACCTCTTCAATAGTACAGTCTTCTGAG 4740  
|||||  
4800 CATGGAGCAGACCTCTTCAACAGCTACACCTCTTCAATAGTACAGTCTTCTGAG 4859  
|||||  
4741 CATGGAGCAGACCTCTTCAACAGCTACACCTCTTCAATAGTACAGTCTTCTGAG 4800  
|||||  
4860 CATGGAGCAGACCTCTTCAACAGCTACACCTCTTCAATAGTACAGTCTTCTGAG 4919  
|||||  
4801 CATGGAGCAGACCTCTTCAACAGCTACACCTCTTCAATAGTACAGTCTTCTGAG 4860  
|||||  
4920 CATGGAGCAGACCTCTTCAACAGCTACACCTCTTCAATAGTACAGTCTTCTGAG 4979  
|||||  
4861 CATGGAGCAGACCTCTTCAACAGCTACACCTCTTCAATAGTACAGTCTTCTGAG 4920  
|||||  
4980 CATGGAGCAGACCTCTTCAACAGCTACACCTCTTCAATAGTACAGTCTTCTGAG 5039  
|||||  
4921 CATGGAGCAGACCTCTTCAACAGCTACACCTCTTCAATAGTACAGTCTTCTGAG 4980  
|||||  
5040 CATGGAGCAGACCTCTTCAACAGCTACACCTCTTCAATAGTACAGTCTTCTGAG 5099  
|||||  
4981 CATGGAGCAGACCTCTTCAACAGCTACACCTCTTCAATAGTACAGTCTTCTGAG 5040  
|||||  
5100 CATGGAGCAGACCTCTTCAACAGCTACACCTCTTCAATAGTACAGTCTTCTGAG 5159  
|||||  
5041 CATGGAGCAGACCTCTTCAACAGCTACACCTCTTCAATAGTACAGTCTTCTGAG 5100  
|||||  
5160 CATGGAGCAGACCTCTTCAACAGCTACACCTCTTCAATAGTACAGTCTTCTGAG 5219  
|||||  
5101 CATGGAGCAGACCTCTTCAACAGCTACACCTCTTCAATAGTACAGTCTTCTGAG 5160  
|||||  
5220 CATGGAGCAGACCTCTTCAACAGCTACACCTCTTCAATAGTACAGTCTTCTGAG 5279  
|||||  
5161 CATGGAGCAGACCTCTTCAACAGCTACACCTCTTCAATAGTACAGTCTTCTGAG 5220  
|||||  
5280 CATGGAGCAGACCTCTTCAACAGCTACACCTCTTCAATAGTACAGTCTTCTGAG 5339  
|||||  
5221 CATGGAGCAGACCTCTTCAACAGCTACACCTCTTCAATAGTACAGTCTTCTGAG 5280  
|||||  
5340 CATGGAGCAGACCTCTTCAACAGCTACACCTCTTCAATAGTACAGTCTTCTGAG 5399  
|||||  
5281 CATGGAGCAGACCTCTTCAACAGCTACACCTCTTCAATAGTACAGTCTTCTGAG 5340  
|||||





QY	1621	CAGAAATGGTCAGATGATGATATTTCTCTAATAGTGGTCATGAGAAATAAACAAAGGTGAT	1680
Db	1740	CAGAAATGGTCAGATGATGATATTTCTAATAGTGGTCATGAGAAATAAACAAAGGTGAT	1799
QY	1681	TCTATTCAGAATGAGAAAAATCCCTAACCCATAGAAATCACTCGAAAAAGATCTGGCTTC	1740
Db	1800	TCTATTCAGAATGAGAAAAATCCCTAACCCATAGAAATCACTCGAAAAAGATCTGGCTTC	1859
QY	1741	AAAACGAAAGCTGAACCTATTAAGCAGCAGTATTAAGCAATATGGAACCTCGAAATTAATATYC	1800
Db	1860	AAAACGAAAGCTGAACCTATTAAGCAGCAGTATTAAGCAATATGGAACCTCGAAATTAATATYC	1919
QY	1801	CACAACTTCAAAAAGCCTTAAAAAGATAGGGCGAGAGAGAGTCTCTACAGCATATYT	1860
Db	1920	CACAACTTCAAAAAGCCTTAAAAAGATAGGGCGAGAGAGAGTCTCTACAGCATATYT	1979
QY	1861	CATGGCGCTTGAACCTAGTAGTCAGTATGAATCTAAGCCCACTAATTTAGTGAATTGCA	1920
Db	1980	CATGGCGCTTGAACCTAGTAGTCAGTATGAATCTAAGCCCACTAATTTAGTGAATTGCA	2039
QY	1921	ATTGATAGTTGTTCTAGCAGTGAAGATTAAGAAAAAAAGTACACCAAAATGCCAGTC	1980
Db	2040	ATTGATAGTTGTTCTAGCAGTGAAGATTAAGAAAAAAAGTACACCAAAATGCCAGTC	2099
QY	1981	AGGCACACAGAAACCTCAACTCACTGTGAAGSTAAAGAACTGCACTGGAGCCAAAGAG	2040
Db	2100	AGGCACACAGAAACCTCAACTCTGTGAAGSTAAAGAACTGCACTGGAGCCAAAGAG	2159
QY	2041	AGTAAACAAGCCAAATGAACAGACAGCAAGTAAAAAGACATGACAGTACTTCCAGAGCTG	2100
Db	2160	AGTAAACAAGCCAAATGAACAGACAGCAAGTAAAAAGACATGACAGTACTTCCAGAGCTG	2219
QY	2101	AAGTTAAACAAATGCACTCGGTTCTTTTACTAAGTGTCTAAATACCAGTAACTTAAAGAA	2160
Db	2220	AAGTTAAACAAATGCACTCGGTTCTTTTACTAAGTGTCTAAATACCAGTAACTTAAAGAA	2279
QY	2161	TTTGTCAATCCGTAGCCTTCCAAAGAGAAACAAAAAGAGAAACTGAACAGTTAAAGTG	2220
Db	2280	TTTGTCAATCCGTAGCCTTCCAAAGAGAAACAAAAAGAGAAACTGAACAGTTAAAGTG	2339
QY	2221	TCTAATAATGCTGAAGACCCCAAGACATGTCATGTTAACTGAGAAAGGGTTTGCAAACT	2280
Db	2340	TCTAATAATGCTGAAGACCCCAAGACATGTCATGTTAACTGAGAAAGGGTTTGCAAACT	2399
QY	2281	GAAAGATCTGTAGAGAGTAGACAGATATTTTCATTTGGTACTGGTACTGATTAAGCAGCTCAG	2340
Db	2400	GAAAGATCTGTAGAGAGTAGACAGATATTTTCATTTGGTACTGGTACTGATTAAGCAGCTCAG	2459
QY	2341	GAAAGATCTGCTTACTGGAAGTTAGCCTCTAGGGAAGGCAAAAACGAAACCAATTAAT	2400
Db	2460	GAAAGATCTGCTTACTGGAAGTTAGCCTCTAGGGAAGGCAAAAACGAAACCAATTAAT	2519
QY	2401	TGTGTGACTAGTGTGCAGCATTTTGAAAAACCCCAAGGCACTAATTCATGTGTTTCCAAA	2460
Db	2520	TGTGTGACTAGTGTGCAGCATTTTGAAAAACCCCAAGGCACTAATTCATGTGTTTCCAAA	2579
QY	2461	GATATATAGAAATGACACAGAAAGGCTTTTAGTATTCATTTGGCAGATGAAGTTAACCAAGT	2520
Db	2580	GATATATAGAAATGACACAGAAAGGCTTTTAGTATTCATTTGGCAGATGAAGTTAACCAAGT	2639
QY	2521	CGGGAAACACATATGGAATGGAAGAAAGTAACTGTGATGCTGAGTATTTTSCAATATACA	2580
Db	2640	CGGGAAACACATATGGAATGGAAGAAAGTAACTGTGATGCTGAGTATTTTSCAATATACA	2699
QY	2581	TTCAAGGTTTCAAAAGCGCCAGCTCATTTTGCTCTGTTTTCAAATCCAGGAAATGCAAGAG	2640
Db	2700	TTCAAGGTTTCAAAAGCGCCAGCTCATTTTGCTCTGTTTTCAAATCCAGGAAATGCAAGAG	2759
QY	2641	GAAATGTGAATATCTCTGCCCACTCTGGGTCTTTAAAGAAACAAAGTCCAAAAGTCACT	2700
Db	2760	GAAATGTGAATATCTCTGCCCACTCTGGGTCTTTAAAGAAACAAAGTCCAAAAGTCACT	2819

QY	2701	TTTTGATGTGAACAAAGAGAAATATCAAGGAAGANTGATCTAAATATACAGCTGT	2760
Db	2820	TTTTGATGTGAACAAAGAGAAATATCAAGGAAGANTGATCTAAATATACAGCTGT	2879
QY	2761	CAGACAGTTAAATATCATCTGCAGAGGCTTCTCGTGGTGTGCAGAAAGATAAAGCCAGTTGAT	2820
Db	2880	CAGACAGTTAAATATCATCTGCAGAGGCTTCTCGTGGTGTGCAGAAAGATAAAGCCAGTTGAT	2939
QY	2821	AATGCCAAATGTAGTATCAAAAGAGGCTCTAGGTTTGTCTATCATCTCAGTTCAGAGGC	2880
Db	2940	AATGCCAAATGTAGTATCAAAAGAGGCTCTAGGTTTGTCTATCATCTCAGTTCAGAGGC	2999
QY	2881	AACGAAATCGACATCTACTCTCCATATAACATGACATGACCTTTTACAAACCCATATCTGTATA	2940
Db	3000	AACGAAATCGACATCTACTCTCCATATAACATGACATGACCTTTTACAAACCCATATCTGTATA	3059
QY	2941	CCACCACCTTTTCCCATCAAGTCATCTTGTATAAATCAATATGAAAGAAATCTGCTAGAG	3000
Db	3060	CCACCACCTTTTCCCATCAAGTCATCTTGTATAAATCAATATGAAAGAAATCTGCTAGAG	3119
QY	3001	GAAGAACTTTGAGGAACATTCATGTCTACCTGAAAGAGAAATGGGAAATGAGAACATTTCCA	3060
Db	3120	GAAGAACTTTGAGGAACATTCATGTCTACCTGAAAGAGAAATGGGAAATGAGAACATTTCCA	3179
QY	3061	AGTACAGTGCAGACACATTAGCCGCTATATACATTATGAGAAATGTTTTTAAAGAACGACG	3120
Db	3180	AGTACAGTGCAGACACATTAGCCGCTATATACATTATGAGAAATGTTTTTAAAGAACGACG	3239
QY	3121	TCACACCATATTTAATGAAGTAGGTTCCAGTACTAATGAAGTGGGCTCCAGTATTTAATGA	3180
Db	3240	TCACACCATATTTAATGAAGTAGGTTCCAGTACTAATGAAGTGGGCTCCAGTATTTAATGA	3299
QY	3181	ATAGGTTCCAGTAGTGAAGAACATTCACAGACAATAGGTAGAGAAACAGAGGCCAAATTTG	3240
Db	3300	ATAGGTTCCAGTAGTGAAGAACATTCACAGACAATAGGTAGAGAAACAGAGGCCAAATTTG	3359
QY	3241	AATGCTATGCTTAGATTTAGGGGTTTTGCAACCTGAGGCTCTATTAACAAGCTTCTCGGA	3300
Db	3360	AATGCTATGCTTAGATTTAGGGGTTTTGCAACCTGAGGCTCTATTAACAAGCTTCTCGGA	3419
QY	3301	AGTAATTTGAAGCACCCTGAAATATAAAGACAGAAATGTGAAGAAAGTAGTTCAACACTGTT	3360
Db	3420	AGTAATTTGAAGCACCCTGAAATATAAAGACAGAAATGTGAAGAAAGTAGTTCAACACTGTT	3479
QY	3361	AATACAGATTTCTCTCATATCTGATTTCAAGATACTTAGAACACGCCATGGAAGTAGT	3420
Db	3480	AATACAGATTTCTCTCATATCTGATTTCAAGATACTTAGAACACGCCATGGAAGTAGT	3539
QY	3421	CATGCATCTCAGGTTTTGTTCTGAGACACCTGATGACCTGTTATGATGATGTGAATAAAG	3480
Db	3540	CATGCATCTCAGGTTTTGTTCTGAGACACCTGATGACCTGTTATGATGATGTGAATAAAG	3599
QY	3481	GAAGTACTAGTTTGGTGAAAAATATCAATTAAGAAAGAACTTCGCTGTTTAGCAAAAGC	3540
Db	3600	GAAGTACTAGTTTGGTGAAAAATATCAATTAAGAAAGAACTTCGCTGTTTAGCAAAAGC	3659
QY	3541	GTCCAGAAAGAGAGCTTAGCAGAGAGTCTTAGCCCTTTCACCCATATACATTTGGCTAG	3600
Db	3660	GTCCAGAAAGAGAGCTTAGCAGAGAGTCTTAGCCCTTTCACCCATATACATTTGGCTAG	3719
QY	3601	GGTTACCGAAGAGGGGCCAAGAATTTAGAGTCTCAGAGAAGAACTTATCTAGTAGGAT	3660
Db	3720	GGTTACCGAAGAGGGGCCAAGAATTTAGAGTCTCAGAGAAGAACTTATCTAGTAGGAT	3779
QY	3661	GAAAGAGCTTCCCTCTTCCAAACACTTGTATTTTGGTAAAGTAAACAATATACCTTCCAG	3720
Db	3780	GAAAGAGCTTCCCTCTTCCAAACACTTGTATTTTGGTAAAGTAAACAATATACCTTCCAG	3839
QY	3721	TCTACTAGAGCTTACACCGTTGCTTACCGAGTGTGTCTCTAAGAACACAGAGAGAAATTTA	3780
Db	3840	TCTACTAGAGCTTACACCGTTGCTTACCGAGTGTGTCTCTAAGAACACAGAGAGAAATTTA	3899
QY	3781	TTATCTATTGAAGAAATAGCTTAATATGACTGCAGTAACAGAGTAATATTGGCAAGGACCTCT	3840







Query Match 99.98; Score 5585.8; DB 6; Length 5712;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 5587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 ATGATTTATCTGCTTCCTGCGTGAAGATACAAATATGTCATTAATGCTATGCGAAA 60  
|||||  
120 ATGATTTATCTGCTTCCTGCGTGAAGATACAAATATGTCATTAATGCTATGCGAAA 179  
61 ATCTAGAGTGTCCCATCTGTCTGAGTGTATCAAGAACCTGTCCCAAAAGTGTGAC 120  
|||||  
180 ATCTAGAGTGTCCCATCTGTCTGAGTGTATCAAGAACCTGTCTCCCAAAAGTGTGAC 239  
121 CACATATTTTGCATAATTTTGCATGCTGAATCTTCTACACCAAGAAAGGCTTCACAG 180  
240 CACATATTTTGCATAATTTTGCATGCTGAATCTTCTACACCAAGAAAGGCTTCACAG 299  
181 TGTCTTATGTAAATGATATTAACCAAGAGGCTTCACCAAGAAAGTTCAGATTTACT 240  
|||||  
300 TGTCTTATGTAAATGATATTAACCAAGAGGCTTCACCAAGAAAGTTCAGATTTACT 359  
241 CAACCTTGTGAAGAGCTATTTGAATATTTGTCTTTCACCTTGACACAGTTTGGAG 300  
|||||  
360 CAACCTTGTGAAGAGCTATTTGAATATTTGTCTTTCACCTTGACACAGTTTGGAG 419  
301 TATGCAAAACAGCTATTAATTTTGCATAATTTTGCATCTCTGACATCTTAAAGAT 360  
|||||  
420 TATGCAAAACAGCTATTAATTTTGCATAATTTTGCATCTCTGACATCTTAAAGAT 479  
361 GAAGTTTCTATCATCCAAAGTATGGGCTACAGAAACCGTCCCAAAAGACTTCTACAGAT 420  
|||||  
480 GAAGTTTCTATCATCCAAAGTATGGGCTACAGAAACCGTCCCAAAAGACTTCTACAGAT 539  
421 GAACCCGAAAAATCCTTCTGCGAGAAACCACTCTAGTGTCCAACTCTCTAACCTTGA 480  
|||||  
540 GAACCCGAAAAATCCTTCTGCGAGAAACCACTCTAGTGTCCAACTCTCTAACCTTGA 599  
481 ACTGTGAGAACTCTGAGACAAAGCAGGATACAACTCAAAAGAGCTCTCTACAT 540  
|||||  
600 ACTGTGAGAACTCTGAGACAAAGCAGGATACAACTCAAAAGAGCTCTCTACAT 659  
541 GAATTTGGATCTGATTTCTTCTGAAATACCGTTAATTAAGGCAACTTAATGCGAGTGGGA 600  
|||||  
660 GAATTTGGATCTGATTTCTTCTGAAATACCGTTAATTAAGGCAACTTAATGCGAGTGGGA 719  
601 GATCAGAATTTGTACAAATCACCCTCAAGGAAACAGGATGAATCAGTTTGGATCT 660  
720 GATCAGAATTTGTACAAATCACCCTCAAGGAAACAGGATGAATCAGTTTGGATCT 779  
661 GCAAAAAAGGCTGCTGTGGAATTTTCTGAGAGGATGTATACAAATCTGACATCTCAA 720  
|||||  
780 GCAAAAAAGGCTGCTGTGGAATTTTCTGAGAGGATGTATACAAATCTGACATCTCAA 839  
721 CCCGATTAATTAATGATTTGAACACCCTGAGAAAGCTGAGGATGAGGATCCAGAAAAG 780  
|||||  
840 CCCGATTAATTAATGATTTGAACACCCTGAGAAAGCTGAGGATGAGGATCCAGAAAAG 899  
781 TATCAGGCTAGTCTGTTTCAAACTTGATGTGAGAGGATGAGGATGAGGATCCATGCTCC 840  
900 TATCAGGCTAGTCTGTTTCAAACTTGATGTGAGAGGATGAGGATGAGGATCCATGCTCC 959  
841 AGCTCATTTACAGCATGAGAACAGAGCTTATTAATCTCACTAAAGACAGATGATAGAA 900  
|||||  
960 AGCTCATTTACAGCATGAGAACAGAGCTTATTAATCTCACTAAAGACAGATGATAGAA 1019  
901 AAGGCTGAATCTGTAATTAAGAAACAGAGCTGAGGATGAGGATGAGGATCCATTAACGA 960  
|||||  
1020 AAGGCTGAATCTGTAATTAAGAAACAGAGCTGAGGATGAGGATGAGGATCCATTAACGA 1079  
961 TGGGCTGGAAGTAAGAAACATGTATGATAGGAGGATCCAGGACAGAAAGAAAGTA 1020  
1080 TGGGCTGGAAGTAAGAAACATGTATGATAGGAGGATCCAGGACAGAAAGAAAGTA 1139  
1021 GATCTGAATGCTGATCCCTGTGTGAGAGAAAGAAAGATGAATTAAGAGAAACTGCTATGC 1080

|||||  
1140 GATCTGAATGCTGATCCCTGTGTGAGAGAAAGAAAGATGAATTAAGCAAAACTGCTATGC 1199  
1081 TCAGAGAAATCTAGAGATCTGAGAAATGCTTCTGATTAACAGTAATTAAGACATTCAG 1140  
1200 TCAGAGAAATCTAGAGATCTGAGAAATGCTTCTGATTAACAGTAATTAAGACATTCAG 1259  
1141 AAGTTAATGAGTGTGTTTCCAGAGTATGAACTGATGAGTGTGATGATGATCATGAT 1200  
1260 AAGTTAATGAGTGTGTTTCCAGAGTATGAACTGATGAGTGTGATGATGATCATGAT 1319  
1201 GGGGAGTCTGAATCAAAATGCCAAATGATGATGATGATGATGATGATGATGATGAT 1260  
1320 GGGGAGTCTGAATCAAAATGCCAAATGATGATGATGATGATGATGATGATGATGAT 1379  
1261 GAATATTTCTGTTCTTCAGAGAAATTAACCTTACGAGGATGATGATGATGATGATGAT 1320  
1380 GAATATTTCTGTTCTTCAGAGAAATTAACCTTACGAGGATGATGATGATGATGATGAT 1439  
1321 ATATGTAAAGTGAAGAGTTCACCTCAAAATCAGTGAAGTATTAATTAAGACAAATA 1380  
1440 ATATGTAAAGTGAAGAGTTCACCTCAAAATCAGTGAAGTATTAATTAAGACAAATA 1499  
1381 TTTGGAAAACTATCGAGAGAGCAAGCCTCCCAACTTAAGCCATGTAAGTGAAT 1440  
1500 TTTGGAAAACTATCGAGAGAGCAAGCCTCCCAACTTAAGCCATGTAAGTGAAT 1559  
1441 CTAATTAATGAGCAATTTTGTACTGAGCAGACAGATTAATACAAAGCGTCCCTACAAAT 1500  
1560 CTAATTAATGAGCAATTTTGTACTGAGCAGACAGATTAATACAAAGCGTCCCTACAAAT 1619  
1501 AAATTAAGCGTAAAGAGAGACCTACATCAGAGCCTTCTCTGAGGATTTTATCAATAA 1560  
1620 AAATTAAGCGTAAAGAGAGACCTACATCAGAGCCTTCTCTGAGGATTTTATCAATAA 1679  
1561 GCAGATTTGGCAATTCAGAAAGCTCTGAAATGATTAATCAGGAGTAACCAACGAG 1620  
1680 GCAGATTTGGCAATTCAGAAAGCTCTGAAATGATTAATCAGGAGTAACCAACGAG 1739  
1621 CAGATGCTCAAGTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1680  
1740 CAGATGCTCAAGTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1799  
1681 TCTATTCAAGATGAGAAAAATCTTAACCAATTAATTAATTAATTAATTAATTAATTAAT 1740  
1800 TCTATTCAAGATGAGAAAAATCTTAACCAATTAATTAATTAATTAATTAATTAATTAAT 1859  
1741 AAACGAAAGCTGAACCTATTAAGCAGCTATTAAGCAATTAATTAATTAATTAATTAAT 1800  
1860 AAACGAAAGCTGAACCTATTAAGCAGCTATTAAGCAATTAATTAATTAATTAATTAAT 1919  
1801 CACATTTCAAAAGCCTTAAGAAAGATAGGCTGAGGAGGAAAGTCTTACCAGGCAATTT 1860  
1920 CACATTTCAAAAGCCTTAAGAAAGATAGGCTGAGGAGGAAAGTCTTACCAGGCAATTT 1979  
1861 CATGGCTTGAAGTATGATGAGTAAATTAAGCCACCTTAATTAATTAATTAATTAATTAAT 1920  
1980 CATGGCTTGAAGTATGATGAGTAAATTAAGCCACCTTAATTAATTAATTAATTAATTAAT 2039  
1921 ATTGATGTTGTTCTAGCAGTGAAGATTAAGAAAAAGTAACAACCAATGCGCAGTC 1980  
2040 ATTGATGTTGTTCTAGCAGTGAAGATTAAGAAAAAGTAACAACCAATGCGCAGTC 2099  
1981 AGGACAGCAGAAACCTTAACCTGATGAGGATTAAGAAACCTGCACTGAGGCAAGAAAG 2040  
2100 AGGACAGCAGAAACCTTAACCTGATGAGGATTAAGAAACCTGCACTGAGGCAAGAAAG 2159  
2041 AGTAACAAGCCAAATGAACAGACAGATTAAGAAAGATGATGATGATGATGATGATGAT 2100  
2160 AGTAACAAGCCAAATGAACAGACAGATTAAGAAAGATGATGATGATGATGATGATGAT 2219  
2101 AAGTTAACAATGCACTGCTTCTTTACTAAGTGTCAAAATACCACTGAATTAAGAA 2160  
|||||

Db 2220 AAGTAAACAATGCACCTGGTCTTTTACTAAGTGTTCATAATACACAGTAACTTAAGAA 2279  
QY 2161 TTGTCAATCCCTGACCTTTCCAGAGAAAGAAAGAAAGAAAGTAAAGTAAAGT 2220  
Db 2280 TTGTCAATCCCTGACCTTTCCAGAGAAAGAAAGAAAGTAAAGTAAAGT 2239  
QY 2221 TCTAATATGCTGAAAGACCCCAAGATCTCATGTTAAGTGAAGAGGTTTTCAGAACT 2280  
Db 2340 TCTAATATGCTGAAAGACCCCAAGATCTCATGTTAAGTGAAGAGGTTTTCAGAACT 2239  
QY 2281 GAAAGATCTGTAGAGAGTACGATATTTCAATGTTGTTACTGTTATGAGCAGT 2340  
Db 2400 GAAAGATCTGTAGAGAGTACGATATTTCAATGTTGTTACTGTTATGAGCAGT 2459  
QY 2341 GAAAGATCTGTAGAGAGTACGATATTTCAATGTTGTTACTGTTATGAGCAGT 2400  
Db 2460 GAAAGATCTGTAGAGAGTACGATATTTCAATGTTGTTACTGTTATGAGCAGT 2519  
QY 2401 TGTGTGAGTCAAGTGTGAGCATTTTGAAGACCCCAAGGAGTAAATTCATGTTTCCAA 2460  
Db 2520 TGTGTGAGTCAAGTGTGAGCATTTTGAAGACCCCAAGGAGTAAATTCATGTTTCCAA 2579  
QY 2461 GATATATGAATGACACAGAAAGCTTTAAGTATCCATTTGGACATGAAGTTAACACAGT 2520  
Db 2580 GATATATGAATGACACAGAAAGCTTTAAGTATCCATTTGGACATGAAGTTAACACAGT 2639  
QY 2521 CGGGAACAAGCATAGAAATGGAAGAAAGTGAAGTATGCTGATGTTTTCAGAAATACA 2580  
Db 2640 CGGGAACAAGCATAGAAATGGAAGAAAGTGAAGTATGCTGATGTTTTCAGAAATACA 2699  
QY 2581 TTCAAGTTTCAAGAGCCGAGTATTTGCTGTTTTCATTCACAGGAAATGAGAGAG 2640  
Db 2700 TTCAAGTTTCAAGAGCCGAGTATTTGCTGTTTTCATTCACAGGAAATGAGAGAG 2759  
QY 2641 GAATGTGCAACATTTCTGCCCCACTGCTGGTCTTAAAGAAACAAAGTCCAAAGTACT 2700  
Db 2760 GAATGTGCAACATTTCTGCCCCACTGCTGGTCTTAAAGAAACAAAGTCCAAAGTACT 2819  
QY 2701 TTGAATGTGAACAAAAGAAAGAAATCAAGAAAGTATGATTAATCAAGCCGTGA 2760  
Db 2820 TTGAATGTGAACAAAAGAAAGAAATCAAGAAAGTATGATTAATCAAGCCGTGA 2879  
QY 2761 CAGACAGTAAATTCACCTGAGGCTTCTGTTGTTGTCAGAAAGTAAAGCAGTTGAT 2820  
Db 2880 CAGACAGTAAATTCACCTGAGGCTTCTGTTGTTGTCAGAAAGTAAAGCAGTTGAT 2939  
QY 2821 AATGCCAAATGTAGTATCAAGAGAGGCTTAGGTTTGTCTATCTCAGTTCAGAGGC 2880  
Db 2940 AATGCCAAATGTAGTATCAAGAGAGGCTTAGGTTTGTCTATCTCAGTTCAGAGGC 2999  
QY 2881 AAGCAAACTGGACTCTTACTCCAAATTAACATGAGCTTTTACAAAACCATATGCTATA 2940  
Db 3000 AAGCAAACTGGACTCTTACTCCAAATTAACATGAGCTTTTACAAAACCATATGCTATA 3059  
QY 2941 CCACACATTTTCCCATCAAGTATTTGTTAAACTAAATTAAGAAAAATCTGCTAGAG 3000  
Db 3060 CCACACATTTTCCCATCAAGTATTTGTTAAACTAAATTAAGAAAAATCTGCTAGAG 3119  
QY 3001 GAAAACTTTGAGAGACATTTAATGTCACTGAAAGAGAAATGGGAAATGAACATTTCCA 3060  
Db 3120 GAAAACTTTGAGAGACATTTAATGTCACTGAAAGAGAAATGGGAAATGAACATTTCCA 3179  
QY 3061 AGTACAGTGAACAATTTAGCCGTAATTAACATTAAGAGAAAAATGTTTAAAGAAAGCAGC 3120  
Db 3180 AGTACAGTGAACAATTTAGCCGTAATTAACATTAAGAGAAAAATGTTTAAAGAAAGCAGC 3239  
QY 3121 TCAAGCAATATTAATGAAGTGGTCCAGTACTAATGAAGTGGCTCCAGTATTAATGAA 3180  
Db 3240 TCAAGCAATATTAATGAAGTGGTCCAGTACTAATGAAGTGGCTCCAGTATTAATGAA 3299  
QY 3181 ATAGGTTCCAGTATGAAGAAACATTTCAAGCAAGAACTAGTGAAGAAAGAGGCCAAAAATTG 3240  
Db 3300 ATAGGTTCCAGTATGAAGAAACATTTCAAGCAAGAACTAGTGAAGAAAGAGGCCAAAAATTG 3359

QY 3241 AATGCTATGCTTAGATTAGGGGTTTGGCAACCTGAGGCTTAATAACAAGTCTTCCGGA 3300  
Db 3360 AATGCTATGCTTAGATTAGGGGTTTGGCAACCTGAGGCTTAATAACAAGTCTTCCGGA 3419  
QY 3301 AGTAAATTTGAAGCATCTCTGAAGTAAAAAGCAGAAATATGAAGAGTTCAGACTGTT 3360  
Db 3420 AGTAAATTTGAAGCATCTCTGAAGTAAAAAGCAGAAATATGAAGAGTTCAGACTGTT 3479  
QY 3361 AATACAGATTTCTCTCATATCTGATTTTCAGATTAATTAACAAGCTTATGGAAGTACT 3420  
Db 3480 AATACAGATTTCTCTCATATCTGATTTTCAGATTAATTAACAAGCTTATGGAAGTACT 3539  
QY 3421 CATGATCTCAGGTTTGTGAGACACCTGATGAGCCTGTTAGTATGTTGTAAGTAAAG 3480  
Db 3540 CATGATCTCAGGTTTGTGAGACACCTGATGAGCCTGTTAGTATGTTGTAAGTAAAG 3599  
QY 3481 GAAGATACATGTTTGTGTAAGTAAATGACATTAAGGAAAGTCTGCTTTTACCAAAAGC 3540  
Db 3600 GAAGATACATGTTTGTGTAAGTAAATGACATTAAGGAAAGTCTGCTTTTACCAAAAGC 3659  
QY 3541 GTCCAGAAAGAGAGCTTACAGAGAGTCTAGCCCTTTCACCATACATATTTGGCTCAG 3600  
Db 3660 GTCCAGAAAGAGAGCTTACAGAGAGTCTAGCCCTTTCACCATACATATTTGGCTCAG 3719  
QY 3601 GGTTCAGAGAGGGGCCAAGAAATTAAGAGTCTCCAGAAAGAACTTAATCTAGAGAT 3660  
Db 3720 GGTTCAGAGAGGGGCCAAGAAATTAAGAGTCTCCAGAAAGAACTTAATCTAGAGAT 3779  
QY 3661 GAAGAGCTTCCCTGCTTCCCAACATTTGTTAATGTAAGTAAACAATATACCTTCAG 3720  
Db 3780 GAAGAGCTTCCCTGCTTCCCAACATTTGTTAATGTAAGTAAACAATATACCTTCAG 3839  
QY 3721 TCTACTAGGATAGACACCTGTTGCTACCGAGTGTCTGTAAAGAACACAGAGAGAAATTA 3780  
Db 3840 TCTACTAGGATAGACACCTGTTGCTACCGAGTGTCTGTAAAGAACACAGAGAGAAATTA 3839  
QY 3781 TTATCATTTGAAGATATGATTAATGATGCTGATGACAGTAAATATTTGGCAAGGCTCT 3840  
Db 3900 TTATCATTTGAAGATATGATTAATGATGCTGATGACAGTAAATATTTGGCAAGGCTCT 3959  
QY 3841 CAGGAACATCACCTTAGTGAAGAAACAAATGTTCTGCTAGTGTGTTTCTTACAGTGC 3900  
Db 3960 CAGGAACATCACCTTAGTGAAGAAACAAATGTTCTGCTAGTGTGTTTCTTACAGTGC 4019  
QY 3901 AGTGAATTTGAAGACTTGAATGCAAAATACAAACACCGAGATCTTCTTGATTTGTTCT 3960  
Db 4020 AGTGAATTTGAAGACTTGAATGCAAAATACAAACACCGAGATCTTCTTGATTTGTTCT 4079  
QY 3961 TCCAAACAAATGAGGCTATGCTGAAAGCCAGGAGTTGTTGATGATGCAAGGAAATTG 4020  
Db 4080 TCCAAACAAATGAGGCTATGCTGAAAGCCAGGAGTTGTTGATGATGCAAGGAAATTG 4139  
QY 4021 GTTTCAGATGATGAAGAAAGAGAGAGCGGCTTGAAGAAATTAATCAAGAAAGCAAGC 4080  
Db 4140 GTTTCAGATGATGAAGAAAGAGAGAGCGGCTTGAAGAAATTAATCAAGAAAGCAAGC 4199  
QY 4081 ATGATTTCAACCTTAGTGAAGCAGCATCTGGGTGTGAGAGTGAAGAAACGCTCTGAA 4140  
Db 4200 ATGATTTCAACCTTAGTGAAGCAGCATCTGGGTGTGAGAGTGAAGAAACGCTCTGAA 4259  
QY 4141 GACTGCTCAGAGGCTATGCTGAGAGTGAAGCAATTTTAAACCTGAGAGAGAGGATACATG 4200  
Db 4260 GACTGCTCAGAGGCTATGCTGAGAGTGAAGCAATTTTAAACCTGAGAGAGAGGATACATG 4319  
QY 4201 CAACATTAACCTGATTAAGCTCCAGAGAAATGCTGAAGTAAAGCTGTGTTAAGAACAG 4260  
Db 4320 CAACATTAACCTGATTAAGCTCCAGAGAAATGCTGAAGTAAAGCTGTGTTAAGAACAG 4379  
QY 4261 CATGGAGCCAGGCTTCTAACACCTTACCTCATATGAAGTACTTCTGCTTGGAG 4320  
Db 4380 CATGGAGCCAGGCTTCTAACACCTTACCTCATATGAAGTACTTCTGCTTGGAG 4439

4321 GACCTCGAATTCAGAACACATCAGAAAAAGCAGTATTACTTCAGAGAAAGT 4380  
4440 GACCTCGAATTCAGAACACATCAGAAAAAGCAGTATTACTTCAGAGAAAGT 4499  
4381 AGTGAATACCCCTTAAGCCAGATCCAGAGGCCCTTCTGCTGACAGTTTGGAGTCT 4440  
4500 AGTGAATACCCCTTAAGCCAGATCCAGAGGCCCTTCTGCTGACAGTTTGGAGTCT 4559  
4441 GCAGATAGTTCTCAGTAAATAAAGAACAGAGATGGAAGATCCCTCTTAA 4500  
4560 GCAGATAGTTCTCAGTAAATAAAGAACAGAGATGGAAGATCCCTCTTAA 4619  
4501 TGCCCATCATTAAGATAGTGTACATCAGAGTTGCTGTGGAGTCTTCAGATAGA 4560  
4620 TGCCCATCATTAAGATAGTGTACATCAGAGTTGCTGTGGAGTCTTCAGATAGA 4679  
4561 AACTACCCATTCAGAGAGATCATTAAGTGTGTGATGTGAGAGCAACAGCTGGA 4620  
4680 AACTACCCATTCAGAGAGATCATTAAGTGTGTGATGTGAGAGCAACAGCTGGA 4739  
4621 GAGTCTGGGCAACAGATTTGAGCAAAACATCTTACTTGCAGAGCAAGATCTAGAGGA 4680  
4740 GAGTCTGGGCAACAGATTTGAGCAAAACATCTTACTTGCAGAGCAAGATCTAGAGGA 4799  
4681 ACCCTTACCTGGAATCTGGAATCAGCTCTCTCTGATGACCTGATCTGATCTTCT 4740  
4800 ACCCTTACCTGGAATCTGGAATCAGCTCTCTCTGATGACCTGATCTGATCTTCT 4859  
4741 GAAGACAGAGCCAGAGATCAGCTGTGTGGACATACATCTTCAACCTGCACTG 4800  
4860 GAAGACAGAGCCAGAGATCAGCTGTGTGGACATACATCTTCAACCTGCACTG 4919  
4801 AAAGTTCCCAATTGAAAGTTGAGAAATCTGCCAGAGTCCAGCTGCTGCTCATACT 4860  
4920 AAAGTTCCCAATTGAAAGTTGAGAAATCTGCCAGAGTCCAGCTGCTGCTCATACT 4979  
4861 GATACCTGCTGGGATATATGCAATGGAAGAAAGTGTGAGCAGGAGAAAGCAATGACA 4920  
4980 GATACCTGCTGGGATATATGCAATGGAAGAAAGTGTGAGCAGGAGAAAGCAATGACA 5039  
4921 GCTTCAACAGAAAGGCTCAACAAAGATGTCATGATGTGTGCTGGCTGACCCAGAA 4980  
5040 GCTTCAACAGAAAGGCTCAACAAAGATGTCATGATGTGTGCTGGCTGACCCAGAA 5099  
4981 GAATTTATGCTGCTGATCAAGTTTGCAGAAAAACACATCATTAACTTAATCTAAT 5040  
5100 GAATTTATGCTGCTGATCAAGTTTGCAGAAAAACACATCATTAACTTAATCTAAT 5159  
5041 ACTGAAGAGACTACTATGTTGTTATGAAACAGATGCTGAGTTGTGTGTAAGGACA 5100  
5160 ACTGAAGAGACTACTATGTTGTTATGAAACAGATGCTGAGTTGTGTGTAAGGACA 5219  
5101 CTGAATATTTCTAGAAATGCGGAGAGAAATGGGTAGTATTCTGGGTGACC 5160  
5220 CTGAATATTTCTAGAAATGCGGAGAGAAATGGGTAGTATTCTGGGTGACC 5279  
5161 CAGTCTATTTAAAGAAAGAAATGCTGAATGAGCATGATTTTGAAGTCAGAGAGATGTG 5220  
5280 CAGTCTATTTAAAGAAAGAAATGCTGAATGAGCATGATTTTGAAGTCAGAGAGATGTG 5339  
5221 CTCAATGGAAGAACACCAAGGTCGAAGCGGAGCAAGATCCAGAGCAAGAAAGTGC 5280  
5340 CTCAATGGAAGAACACCAAGGTCGAAGCGGAGCAAGATCCAGAGCAAGAAAGTGC 5399  
5281 TTCAGAGGAGCTAGAAATCTGTTGCTATGGGCTTACCAACATGCCCAGATCTCAAGT 5340  
5400 TTCAGAGGAGCTAGAAATCTGTTGCTATGGGCTTACCAACATGCCCAGATCTCAAGT 5459  
5341 GAATGATGATACAGCTGTGTGCTTGTGTGTAAGAGGCTTTCAATCATTCACCTT 5400  
5460 GAATGATGATACAGCTGTGTGCTTGTGTGTAAGAGGCTTTCAATCATTCACCTT 5519  
5401 GGCACAGGTTGTCACCCAAATTTGTTGTTGTGACAGCCAGATGCTGACAGAGCAATGGC 5460

5520 GGCACAGGTTGTCACCCAAATTTGTTGTGACAGCCAGATGCTTCAGAGAGCAATGGC 5579  
5461 TTCATATGCAATTTGGGAGATGTGTGAGGACCCGTGTGTGACCCGAGAGGGTGTGAC 5520  
5580 TTCATATGCAATTTGGGAGATGTGTGAGGACCCGTGTGTGACCCGAGAGGGTGTGAC 5639  
5521 AGTGTAGCACTTACAGAGTCCAGAGAGTGTGACACCTTACCTGATACCCAGATCCCCAC 5580  
5640 AGTGTAGCACTTACAGAGTCCAGAGAGTGTGACACCTTACCTGATACCCAGATCCCCAC 5699  
5581 AGCCACTAC 5589  
5700 AGCCACTAC 5708  
RESULT 6  
ARL18507 5712 bp DNA linear PAT 16-MAY-2001  
LOCUS ARL18507 Sequence 1 from patent US 6149903.  
DEFINITION ARL18507  
ACCESSION ARL18507.1 GI:14100417  
VERSION ARL18507.1  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 5712)  
AUTHORS Holt, J.T., Jensen, R.A., King, M.-C., Page, D.L., Szabo, C.I.,  
Jettion, T.L., Robinson-Benion, C.L., and Thompson, M.E.,  
Characterized BRCA1 and BRCA2 proteins and screening and  
therapeutic methods based on characterized BRCA1 and BRCA2 proteins  
Patent: US 6149903-A 1 21-NOV-2000;  
JOURNAL Location/Qualifiers  
FEATURES  
source 1..5712  
BASE COUNT 1956 a 1099 c 1274 g 1383 t  
ORIGIN  
Query Match 99.9%; Score 5585.8; DB 6; Length 5712;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 5587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
1 ATGGATTTATCTGCTCTTCCGCTTGAAGAGTACAAATGTCATTAATGCTATGACAGAA 60  
120 ATGGATTTATCTGCTCTTCCGCTTGAAGAGTACAAATGTCATTAATGCTATGACAGAA 179  
61 ATCTTAGAGTGTCCATCTGTGTGAGTGTGATCAAGAACTGTCTCCACAAAGTGTAC 239  
180 ATCTTAGAGTGTCCATCTGTGTGAGTGTGATCAAGAACTGTCTCCACAAAGTGTAC 239  
121 CACATATTTTGCAAATTTTGCATGCTGGAACCTTCAACAGAGAAAGGAGGCTTACAG 180  
240 CACATATTTTGCAAATTTTGCATGCTGGAACCTTCAACAGAGAAAGGAGGCTTACAG 299  
181 TGTCTTTATGTAAGATGATATACCAAGAGGCTTCAAGAAAGTACGAGATTAGT 240  
300 TGTCTTTATGTAAGATGATATACCAAGAGGCTTCAAGAAAGTACGAGATTAGT 359  
241 CAACCTGTGAAGAGTATGAAATCATTTGTGCTTTCAGCTTGCAGACAGGTTGGAG 300  
360 CAACCTGTGAAGAGTATGAAATCATTTGTGCTTTCAGCTTGCAGACAGGTTGGAG 419  
301 TATGCAACAGCTATTAATTTTCAAAAAGGAAATTAATCTCTCGAATCTTAAAGAT 360  
420 TATGCAACAGCTATTAATTTTCAAAAAGGAAATTAATCTCTCGAATCTTAAAGAT 479  
361 GAAGTTTCTATCATCCAAAGTATGGGCTACAGAAACGCTGCCAAAAGACTTCAAGAGT 420  
480 GAAGTTTCTATCATCCAAAGTATGGGCTACAGAAACGCTGCCAAAAGACTTCAAGAGT 539  
421 GAACCCGAAATCTTCTCTGAGAGAAACATCTCATATGTCMACTCTCAACCTTGA 480  
540 GAACCCGAAATCTTCTCTGAGAGAAACATCTCATATGTCMACTCTCAACCTTGA 599

OY	481	ACGTGTGAACCTCGAGAGACAAACACAGGATATCAACCTCAAAGACGTCTGTACTAT	540
OY	482	ACGTGTGAACCTCGAGAGACAAACACAGGATATCAACCTCAAAGACGTCTGTACTAT	540
Db	600	ACTGTGAAACTCTGAGAGACAAACACAGGATATCAACCTCAAAGACGTCTGTACTAT	659
OY	541	GAATTTGGATCTGATTCCTCTGGAAGATACGCTTTATATAGGCAACTTATTTGACGTGGGA	600
Db	660	GAATTTGGATCTGATTCCTCTGGAAGATACGCTTTATATAGGCAACTTATTTGACGTGGGA	719
OY	601	GATCAAGAATTTGTTACAAATCACCCCTCAGGAACACAGGATGAATTCAGTTTGGATTC	660
Db	720	GATCAAGAATTTGTTACAAATCACCCCTCAGGAACACAGGATGAATTCAGTTTGGATTC	779
OY	661	GCAAAAAAGCGTGTGTGAATTTTGTGAGACGGATGTATACAAATPACGAACATCATCA	720
Db	780	GCAAAAAAGCGTGTGTGAATTTTGTGAGACGGATGTATACAAATPACGAACATCATCA	839
OY	721	CCGAGTAAATAATGATTTGAAACACCACCTGAGAAAGCGTGCAGCTGAGAGCATCCAGAAAG	780
Db	840	CCGAGTAAATAATGATTTGAAACACCACCTGAGAAAGCGTGCAGCTGAGAGCATCCAGAAAG	899
OY	781	TATAGGTAATCTCTGTTTTCAAACTTGTGATGTGGAGCATGTGGCACAATATCTATGCC	840
Db	900	TATAGGTAATCTCTGTTTTCAAACTTGTGATGTGGAGCATGTGGCACAATATCTATGCC	959
OY	841	AGCTCATATACAGCATGTAGAACAGCAGTTTATCTCACTAATAAGACAGATTAATGTAGAA	900
Db	960	AGCTCATATACAGCATGTAGAACAGCAGTTTATCTCACTAATAAGACAGATTAATGTAGAA	1019
OY	901	AAGCGTGAATTTCTGTATATAAAGCAAAACAGCTGCGCTTAGCAGAGAGCCAAATPACAGA	960
Db	1020	AAGCGTGAATTTCTGTATATAAAGCAAAACAGCTGCGCTTAGCAGAGAGCCAAATPACAGA	1079
OY	961	TGGCGTGAATTAAGAGAAACATGTATATGATGTGGCGGACTCCACACACAGAAAAAGSTA	1020
Db	1080	TGGCGTGAATTAAGAGAAACATGTATATGATGTGGCGGACTCCACACACAGAAAAAGSTA	1139
OY	1021	GATCTGAATGCTGATCCCTGTGTGAGAGAAAGAATGGAATTAAGACAGAAACTGCCATGC	1080
Db	1140	GATCTGAATGCTGATCCCTGTGTGAGAGAAAGAATGGAATTAAGACAGAAACTGCCATGC	1199
OY	1081	TCACAGAAATCCTTAAGATACTGGAAGATGTTCTTGATTAACATTAATATAGAGATTCAG	1140
Db	1200	TCACAGAAATCCTTAAGATACTGGAAGATGTTCTTGATTAACATTAATATAGAGATTCAG	1259
OY	1141	AAAGTATATGATGCTTTTCCAGAAAGTATGACATCTGTAGCTTGTGATGATCACTCATGAT	1200
Db	1260	AAAGTATATGATGCTTTTCCAGAAAGTATGACATCTGTAGCTTGTGATGATCACTCATGAT	1319
OY	1201	GGGAGATCTGAATCAAAATGCCAAAGTACGTGATGATTTGACGTTCTTAATGAGGTAGAT	1260
Db	1320	GGGAGATCTGAATCAAAATGCCAAAGTACGTGATGATTTGACGTTCTTAATGAGGTAGAT	1379
OY	1261	GAATATTCCTGCTTCTTCAGAGAAATATGACTTATGGCCATGTGATNCCATGAGGCTTTA	1320
Db	1380	GAATATTCCTGCTTCTTCAGAGAAATATGACTTATGGCCATGTGATNCCATGAGGCTTTA	1439
OY	1321	ATATGTAAAGTGAAGAGATTCACTCCAAATCAGTAGAGATTAATTTGAAGACAAATA	1380
Db	1440	ATATGTAAAGTGAAGAGATTCACTCCAAATCAGTAGAGATTAATTTGAAGACAAATA	1499
OY	1381	TTTGGGAAAACTTATGGGAAGAGCAAGCCTCCCAACTTAACCCCATGTATACGTGAAT	1440
Db	1500	TTTGGGAAAACTTATGGGAAGAGCAAGCCTCCCAACTTAACCCCATGTATACGTGAAT	1559
OY	1441	CTAATTTATAGAGCATTTGTTTACTGAGCCACAGATTAATACAAAGAGCGTCCCTTCACAAAT	1500
Db	1560	CTAATTTATAGAGCATTTGTTTACTGAGCCACAGATTAATACAAAGAGCGTCCCTTCACAAAT	1619
OY	1501	AAATTTAAAGCTTAAAGAGAGACCTACATCAGGCTTATCTGTAGGATTTTATCAAGAA	1560
Db	1620	AAATTTAAAGCTTAAAGAGAGACCTACATCAGGCTTATCTGTAGGATTTTATCAAGAA	1679

QY		1561	GAGTTGGCAGTTCAAAAGACCTCGTGAATGTAAATACAGGGACTAATCAAAACGGAG	1620
Dd		1680	GCACAATTGGCAGTTCAAAGACCTCTGAAAATGATAATACAGGGAACTAACCAAAACGGAG	1739
QY		1621	CAGAATGGTCAAAGTAGAATATATTACTAATAGTGTCATCGAGAAATAAACAAGAGTGAT	1680
Dd		1740	CAGAATGGTCAAAGTAGAATATATTACTAATATAGTGTCATCGAGAAATAAACAAGAGTGAT	1799
QY		1681	TCTATTGAGAAATGAGAAAAATCTTAACCCAAATAGATCTACTCGAAAAAGATCTGCTTTC	1740
Dd		1800	TCTATTGAGAAATGAGAAAAATCTTAACCCAAATAGAAATCACCTCGAAAAAGATCTGCTTTC	1859
QY		1741	AAAAAGAAAGCTGAGACTATATAGCAGCATATAGCAATATGGAATCTGAAATTTAAATATC	1800
Dd		1860	AAAAAGAAAGCTGAGACTATATAGCAGCATATAGCAATATGGAATCTGAAATTTAAATATC	1919
QY		1801	CACAATTCAAAAGACCTTAATAAAGGTGAGGAGAAAGTTCTTACCAGGCATATT	1860
Dd		1920	CACAATTCAAAAGACCTTAATAAAGGTGAGGAGAAAGTTCTTACCAGGCATATT	1979
QY		1861	CATGCGCTTGAACTAGTAGTACAGTAAGAAATCTAAGCCACCTAAATTTAGTGAATTCGAA	1920
Dd		1980	CATGCGCTTGAACTAGTAGTACAGTAAGAAATCTAAGCCACCTAAATTTAGTGAATTCGAA	2039
QY		1921	ATTGATATAGTTGTTTACGACATGGAAGACATTAAGAAAAAAAGTTCAAACCAATGCCAGTC	1980
Dd		2040	ATTGATATAGTTGTTTACGACATGGAAGACATTAAGAAAAAAAGTTCAAACCAATGCCAGTC	2099
QY		1981	AGGCACAGCGAAACCTTACAACTGATGGAAGGTAAACCTGCACTGAGCGCCACAG	2040
Dd		2100	AGGCACAGCGAAACCTTACAACTGATGGAAGGTAAACCTGCACTGAGCGCCACAG	2159
QY		2041	AGTAAACAAGCCAAATGGAACAGACAGTAAGAAAGACATGACAGTGTACTTTCCAGAGCTG	2100
Dd		2160	AGTAAACAAGCCAAATGGAACAGACAGTAAGAAAGACATGACAGCGTACTTTCCAGAGCTG	2219
QY		2101	AAGTTAAACAATGACGACTGTTCTTTTACTATAGTGTCCAANTACCACTGAATTTAAAGAA	2160
Dd		2220	AAGTTAAACAATGACGACTGTTCTTTTACTATAGTGTCCAANTACCACTGAATTTAAAGAA	2279
QY		2161	TTTTTCAATCCTACCTCTTCAAGGAAGAAAGAAACAGAAACTAGAAACAGTTAAAGTG	2220
Dd		2280	TTTTTCAATCCTACCTCTTCAAGGAAGAAAGAAACAGAAACTAGAAACAGTTAAAGTG	2339
QY		2221	TCTAATATATGCTGAGAAGCCCCCAAGATCTCATTTAAGTGGAGAAAGGTTTTCACAACT	2280
Dd		2340	TCTAATATATGCTGAGAAGCCCCCAAGATCTCATTTAAGTGGAGAAAGGTTTTCACAACT	2399
QY		2281	GAAAGATCTGTAGAGATAGCAGTATTTTCATTTGCTAGCTGGTACTGATTTAGGCACATCAG	2340
Dd		2400	GAAAGATCTGTAGAGATAGCAGTATTTTCATTTGCTAGCTGGTACTGATTTAGGCACATCAG	2459
QY		2341	GAAAGTATCTTCCTTACTGGAAGTTAGACATCTTAGGGAAGGCCMAAAACAGAACCAATATAA	2400
Dd		2460	GAAAGTATCTTCCTTACTGGAAGTTAGACATCTTAGGGAAGGCCMAAAACAGAACCAATATAA	2519
QY		2401	TGTTGATGATGAGTGTGACGATTTGAAACCCCACAGGACATTAATTCATGTGTTTCCAAA	2460
Dd		2520	TGTTGATGATGAGTGTGACGATTTGAAACCCCACAGGACATTAATTCATGTGTTTCCAAA	2579
QY		2461	GATTAATGAAATGACACAGAAAGGCTTTTAAGTATTCATTTGGAGACATGAAGTTAAACACAGT	2520
Dd		2580	GATTAATGAAATGACACAGAAAGGCTTTTAAGTATTCATTTGGAGACATGAAGTTAAACACAGT	2639
QY		2521	CGGGAACAAGCATTAGAAATGGAAGAAAGTGAACCTGATCTCAGTATTTTGCAAGAAATACA	2580
Dd		2640	CGGGAACAAGCATTAGAAATGGAAGAAAGTGAACCTGATCTCAGTATTTTGCAAGAAATACA	2699
QY		2581	TTTCAAGGTTTTCAAAGCCGACATGATTTGCTGTTTTCATTCACAGGAATTCAGAGAAG	2640
Dd		2700	TTTCAAGGTTTTCAAAGCCGACATGATTTGCTGCGTTTTCATTCACAGGAATTCAGAGAAG	2759
QY		2641	GAATGTGCAACATTTCTTGCCCACTCGGGTCTCTTAAGAACAAGTCCAAAAATGCTACT	2700

Db	2760	GAATGTGCAACATTTCTTGCCCACTCTGGTCTTTAAAGAACAAAGTCCAAAAGTCACT	281.9
OY	2701	TTTGAATGTGAACAAAGAAAGAAATACAGAAAGAAATGAGTCTAATATCAAGCTGTGA	2760
Db	2820	TTTGAATGTGAACAAAGAAAGAAATACAGAAAGAAATGAGTCTAATATCAAGCTGTGA	287.9
OY	2761	CAGACAGTTAATATCATCTGCAAGGCTTTCCTGTGTTGGTCAAGAAAGATAGCCAGTTGAT	2820
Db	2880	CAGACAGTTAATATCATCTGCAAGGCTTTCCTGTGTTGGTCAAGAAAGATAGCCAGTTGAT	293.9
OY	2821	AATGCCAATGTAGATATCAAGAGAGGCTCTGAGTTTGTCTATCATCTCACTTCAGAGGC	2880
Db	2940	AATGCCAATGTAGATATCAAGAGAGGCTCTGAGTTTGTCTATCATCTCACTTCAGAGGC	299.9
OY	2881	AACGAACTGAGACTCATTTACTCCAATTAACATGGACTTTTACAAAACCCATATCGTATA	2940
Db	3000	AACGAACTGAGACTCATTTACTCCAATTAACATGGACTTTTACAAAACCCATATCGTATA	305.9
OY	2941	CCACCACTTTTCCCATCAAGTCAATTTGTTTAAACTTAATGTAAAGAAAATCTGCTAGAG	3000
Db	3060	CCACCACTTTTCCCATCAAGTCAATTTGTTTAAACTTAATGTAAAGAAAATCTGCTAGAG	311.9
OY	3001	GAAAACTTTGGAGAACATTCATATGTCACCTGAAAAGAGAAATGGGAAATGACAATTCGA	3060
Db	3120	GAAAACTTTGGAGAACATTCATATGTCACCTGAAAAGAGAAATGGGAAATGACAATTCGA	317.9
OY	3061	AGTACAGTGAACACAAATTAGCCGTATATACATTATAGAGAAAATGTTTTTAAAGAACCCAGC	3120
Db	3180	AGTACAGTGAACACAAATTAGCCGTATATACATTATAGAGAAAATGTTTTTAAAGAACCCAGC	323.9
OY	3121	TCAMGCAATATTAATGAGTAGTTCAGTACTAATGAATGGGCTCCAGTATTAAATGA	3180
Db	3240	TCAMGCAATATTAATGAGTAGTTCAGTACTAATGAATGGGCTCCAGTATTAAATGA	329.9
OY	3181	ATAGGTCACAGTGAATGAACATTCACACACAACTAGGTAAACACAGAGGCCCAAAATTG	3240
Db	3300	ATAGGTCACAGTGAATGAACATTCACACACAACTAGGTAAACACAGAGGCCCAAAATTG	335.9
OY	3241	AATGCTATGCTTGAATTAAGGGGTTTTGGCACTGAGGCTATTAACAAAGCTTCTCTGGA	3300
Db	3360	AATGCTATGCTTGAATTAAGGGGTTTTGGCACTGAGGCTATTAACAAAGCTTCTCTGGA	341.9
OY	3301	AGTATTTGTAAGCATCTCTGAATATAAAAAGCAAGATATGAAGAATAGTTCAACATGTT	3360
Db	3420	AGTATTTGTAAGCATCTCTGAATATAAAAAGCAAGATATGAAGAATAGTTCAACATGTT	347.9
OY	3361	AATACAGATTTCTCTCCATATCTGATTTCAATTAACATTAGAACAGCCTATGGAAGTAGT	3420
Db	3480	AATACAGATTTCTCTCCATATCTGATTTCAATTAACATTAGAACAGCCTATGGAAGTAGT	353.9
OY	3421	CATGCATCTCAGGTTTTGTTCTGAGACACCTATACCTAGCTAGATGATGGTGAATTAAG	3480
Db	3540	CATGCATCTCAGGTTTTGTTCTGAGACACCTATACCTAGCTAGATGATGGTGAATTAAG	359.9
OY	3481	GAAGATACTAGTTTGTGCTGAATAATGACATTAAAGAAAGTTCTGCTGTTTTAGCAAAAGC	3540
Db	3600	GAAGATACTAGTTTGTGCTGAATAATGACATTAAAGAAAGTTCTGCTGTTTTAGCAAAAGC	365.9
OY	3541	GTCACAGAAAGAGAGCTTAGCAGAGAGTCTTAGCCCTTTCACCCATACACATTTGGCTCAG	3600
Db	3660	GTCACAGAAAGAGAGCTTAGCAGAGAGTCTTAGCCCTTTCACCCATACACATTTGGCTCAG	371.9
OY	3601	GGTTACCGAAGAGGGGCCAAGAAATTAAGATTCAGATCCAGAGAGAACTATCATAGAGAGAT	3660
Db	3720	GGTTACCGAAGAGGGGCCAAGAAATTAAGATTCAGATTCAGAGAGAACTATCATAGAGAGAT	377.9
OY	3661	GAAGAGCTTCCCTGCTTCCAACTGTTTATTTGGTAAAGTAAGAAATATATCTTCTCAG	3720
Db	3780	GAAGAGCTTCCCTGCTTCCAACTGTTTATTTGGTAAAGTAAGAAATATATCTTCTCAG	383.9
OY	3721	TCTACTAGCATAGCACCCGTTGCTACCGAGTGTCTGTTAAAGACACAGAGAGAAATTTA	3780

Db	3840	TCCTACAGGCATAGCACCCTGGCTCTACCCAGTGTCTGTCTTAAGAACACAGAGAGAAATTTA	3859
Qy	3781	TTATATTGAAGAAATAGCTTTAAATGACGTGCAGTAACACAGGTAAATATGGCAAGGACATCT	3840
Db	3900	TTATCATTTGAAGAAATAGCTTTAAATGACGTGCAGTAACACAGGTAAATATGGCAAGGACATCT	3959
Qy	3841	CAGAACATTCACCTTACGTGACGAAACAAATGTTCTGCTACTGTGTTCTTCCACAGTC	3900
Db	3960	CAGGAACATTCACCTTACGTGACGAAACAAATGTTCTGCTACTGTGTTCTTCCACAGTC	4019
Qy	3901	AGTGAATTTGGAAGACTTACTGTGCAAAATCAAAACCCAGATGCCCTTCTGATTTGCTTCT	3960
Db	4020	ACTGATTTGGAAGACTTACTGTGCAAAATCAAAACCCAGATGCCCTTCTGATTTGCTTCT	4079
Qy	3961	TCCAACAATAGAGGCATCACTCTGAAAGCCAGGAGTGTGCTGAGTACAGAAATTG	4020
Db	4080	TCCAACAATAGAGGCATCACTCTGAAAGCCAGGAGTGTGCTGAGTACAGAAATTG	4139
Qy	4021	GTTTCAGATGATGAAGAAAGAGAACCGGGCTTTGGAAAGAAATTAATCAGAAGACCAAGC	4080
Db	4140	GTTTCAGATGATGAAGAAAGAGAACCGGGCTTTGGAAAGAAATTAATCAGAAGACCAAGC	4199
Qy	4081	ATGATTTAAACTTATGTTGTTGAAGACATCTGGGTGTGAGTGTGAACCAAGCGTCTGAA	4140
Db	4200	ATGATTTAAACTTATGTTGTTGAAGACATCTGGGTGTGAGTGTGAACCAAGCGTCTGAA	4259
Qy	4141	GACTGCTCAGGGCTATCCTCTCAGAGTGACATTTTAAACCACTCAGCAGAGGGATACCATG	4200
Db	4260	GACTGCTCAGGGCTATCCTCTCAGAGTGACATTTTAAACCACTCAGCAGAGGGATACCATG	4319
Qy	4201	CACATTAACCTTGATTAACCTCCAGCAGAAATGGCTGACTAGAAAGCTGTGTGAAACAG	4260
Db	4320	CACATTAACCTTGATTAACCTCCAGCAGAAATGGCTGACTAGAAAGCTGTGTGAAACAG	4379
Qy	4261	CATGGAGACCCAGCCCTTCAACAGCTACCCCTTCCATCAATAGTGACTCTTCGCCCTTGAG	4320
Db	4380	CATGGAGACCCAGCCCTTCAACAGCTACCCCTTCCATCAATAGTGACTCTTCGCCCTTGAG	4439
Qy	4321	GACCTGCGAAATCCAGAACAAAGCACATCAGAAAAAGCAGTATTAACTTCACAGAAAAGT	4380
Db	4440	GACCTGCGAAATCCAGAACAAAGCACATCAGAAAAAGCAGTATTAACTTCACAGAAAAGT	4499
Qy	4381	AGTGAATACCTTATPAGCCAGAAATCCAGAAGGCTTTTGTCTGACAAAGTTTGAGGTGCT	4440
Db	4500	AGTGAATACCTTATPAGCCAGAAATCCAGAAGGCTTTTGTCTGACAAAGTTTGAGGTGCT	4559
Qy	4441	GCAGATAGTCTTACCAAGTAAAAAATPAAAGACAGAGTGGAAAGTCAATCCCTTTCTAAA	4500
Db	4560	GCAGATAGTCTTACCAAGTAAAAAATPAAAGACAGAGTGGAAAGTCAATCCCTTTCTAAA	4619
Qy	4501	TGCCCATCATTAGATGATAGGTGTTACATGACAACTTGCTGGGAGTCTTCAGAAATAGA	4560
Db	4620	TGCCCATCATTAGATGATAGGTGTTACATGACAACTTGCTGGGAGTCTTCAGAAATAGA	4679
Qy	4561	AACTAACCATCTCAAGAGAGCTCATTTAAGGTTGTGATGTGGAGAGCAACAGCTGGAA	4620
Db	4680	AACTAACCATCTCAAGAGAGCTCATTTAAGGTTGTGATGTGGAGAGCAACAGCTGGAA	4739
Qy	4621	GAGTGTGGGGCCACAGATTTGACGGAACAATCTTACTTCCCAAGGCAAGATCTAGAGGA	4680
Db	4740	GAGTGTGGGGCCACAGATTTGACGGAACAATCTTACTTCCCAAGGCAAGATCTAGAGGA	4799
Qy	4681	ACCCCTTACCTGGAATCTGGAATAGACCTCTTCTCTGATGACCTTGAAATCTGATCTTCT	4740
Db	4800	ACCCCTTACCTGGAATCTGGAATAGACCTCTTCTCTGATGACCTTGAAATCTGATCTTCT	4859
Qy	4741	GAAACAGAGCCCCAGAGTCAAGCTCAGTTCGTTGGCAACAATACCATTTCAACCTCGCATTG	4800
Db	4860	GAAACAGAGCCCCAGAGTCAAGCTCAGTTCGTTGGCAACAATACCATTTCAACCTCGCATTG	4919
Qy	4801	AAAGTTCCTCCCAATTTGAAGTGTGACGAATCTGCCAGAGTCCAGTCTGCTCATACTACT	4860
Db	4920	AAAGTTCCTCCCAATTTGAAGTGTGACGAATCTGCCAGAGTCCAGTCTGCTCATACTACT	4979

QY	4861	GATACCTCGGGTATTAATGCAATGGAAGAAAGTGTGACACAGGAGAAACCCAGAAATTGCA	4920
Db	4880	GATACCTCGGGTATTAATGCAATGGAAGAAAGTGTGACACAGGAGAAACCCAGAAATTGCA	5039
QY	4921	GCTTCAACAGAAAGGCTCACAAAAGAAATGTCATGCTGTCTGTGGCTGACCCAGAA	4980
Db	5040	GCTTCAACAGAAAGGCTCACAAAAGAAATGTCATGCTGTCTGTGGCTGACCCAGAA	5099
QY	4981	GAATTATTCGTCCTGTGTACAAAGTTTGCCAGAAAACACCACTACTTTAACTAACTAAT	5040
Db	5100	GAATTATTCGTCCTGTGTACAAAGTTTGCCAGAAAACACCACTACTTTAACTAACTAAT	5159
QY	5041	ACTGAAGAAGACTACTCATGTTGTTATGAAAACACATGCTGAGTTTGTGTGAACGGACA	5100
Db	5160	ACTGAAGAAGACTACTCATGTTGTTATGAAAACACATGCTGAGTTTGTGTGAACGGACA	5219
QY	5101	CTGAAATATTTTCTAGGAATTCGCGGAGGAGAAAATGGTAGTTAGCTAATTTCTGGGTACC	5160
Db	5220	CTGAAATATTTTCTAGGAATTCGCGGAGGAGAAAATGGTAGTTAGCTAATTTCTGGGTACC	5279
QY	5161	CAGCTCTATTAAGAAAGAAAATGCTGAATAGCATGATTTTGAAGTCAGAGAGATGTG	5220
Db	5280	CAGCTCTATTAAGAAAGAAAATGCTGAATAGCATGATTTTGAAGTCAGAGAGATGTG	5339
QY	5221	GTCATATGGAAGAAACACCAAGCTCCAAAGGAGACAAGAAATCCAGAGACAGAAAGATC	5280
Db	5340	GTCATATGGAAGAAACACCAAGCTCCAAAGGAGACAAGAAATCCAGAGACAGAAAGATC	5399
QY	5281	TTCAAGGGGGCTAGAAATCTGTTGCTATAGGGCCCTTCACCAACATGCCACAGATCAACTG	5340
Db	5400	TTCAAGGGGGCTAGAAATCTGTTGCTATAGGGCCCTTCACCAACATGCCACAGATCAACTG	5459
QY	5341	GAATGGAATGGTACACGCTGTGTGTGCTCTGTGTGTGTAAGAGCTTTATCATTCACCTT	5400
Db	5460	GAATGGAATGGTACACGCTGTGTGTGCTCTGTGTGTGTAAGAGCTTTATCATTCACCTT	5519
QY	5401	GGCACAGGTGCCACCCCAATTTGTGTGTGTGCAGCCAGATGCTGTGCAGACAGAGACAATGC	5460
Db	5520	GGCACAGGTGCCACCCCAATTTGTGTGTGTGCAGCCAGATGCTGTGCAGACAGAGACAATGC	5579
QY	5461	TTCCATGCAATTTGGGCGAGATGTGTGAGGACACCTGTGTGACCCGAGAGTGGTGTGGAC	5520
Db	5580	TTCCATGCAATTTGGGCGAGATGTGTGAGGACACCTGTGTGACCCGAGAGTGGTGTGGAC	5639
QY	5521	AGTGTAGACACTCTACCAAGTGTGCAGAGAGTGGACACCACTCATCGATACCCAGATCCCCAC	5580
Db	5640	AGTGTAGACACTCTACCAAGTGTGCAGAGAGTGGACACCACTCATCGATACCCAGATCCCCAC	5699
QY	5581	AGCCACTAC 5589	
Db	5700	AGCCACTAC 5708	
RESULT 7			
ARI125601			
LOCUS	ARI125601	5712 bp	DNA
DEFINITION	Sequence 1 from patent US 6177410.		linear
ACCESSION	ARI25601		
VERSION	ARI25601.1		
KEYWORDS	GI:14111663		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 5712):		
AUTHORS	Holt,J.T., Jensen,R.A., King,M.-C., Steiner,M.S.,		
TITLE	Rodinson-Benion,C.L. and Thompson,M.E.		
JOURNAL	Therapeutic methods for prostate cancer		
FEATURES	Patent: US 6177410-A 1 23-JAN-2001;		
source	Location/Qualifiers		
	1..5712		
	/organism="unknown"		
BASE COUNT	1956 a 1099 c 1274 g 1383 t		

Query Match.	99.9%	Score 5585.8	DB 6	Length 5712
Best Local Similarity	100.0%	Pred. No. 0		
Matches 5587	Conservative 0	Mismatches 2	Indels 0	Gaps 0
OY	1	ATGATATTATCGCTCGGCGTTGTAAGACAGTAAAGTAAATGTCATTAATGCTATGCGAAA	60	
Db	120	ATGATATTATCGCTCGGCGTTGTAAGACAGTAAAGTAAATGTCATTAATGCTATGCGAAA	179	
OY	61	ATCTTAGAGTGTCCCATCTGTGTGAGTTGATCAAGAACCTGTCTCCCAAAAGTGTGAC	120	
Db	180	ATCTTAGAGTGTCCCATCTGTGTGAGTTGATCAAGAACCTGTCTCCCAAAAGTGTGAC	239	
OY	121	CACATATTTTCCAAATTTTGCATGTGAACCTCTCAACAGAAAGGCGCTTCACAG	180	
Db	240	CACATATTTTCCAAATTTTGCATGTGAACCTCTCAACAGAAAGGCGCTTCACAG	299	
OY	181	TGTCCTTATGTAAAGATATTAACCAAAAGAGCCCTACAAAGAAAGTACGAGATTAGT	240	
Db	300	TGTCCTTATGTAAAGATATTAACCAAAAGAGCCCTACAAAGAAAGTACGAGATTAGT	359	
OY	241	CAACTTGTGAAGAGCTATTGAAAATCATTTTGTGCTTTCAGCTTGACAGAGTTTGGAG	300	
Db	360	CAACTTGTGAAGAGCTATTGAAAATCATTTTGTGCTTTCAGCTTGACAGAGTTTGGAG	419	
OY	301	TATGCAAAACGCTATTAATTTTGCAAAAAAGGAAAATACTCTCTGAACATCTAAAGAT	360	
Db	420	TATGCAAAACGCTATTAATTTTGCAAAAAAGGAAAATACTCTCTGAACATCTAAAGAT	479	
OY	361	GAACTTTCTATCATCCAAAGTATGGGCTACAGAAACCGTGGCAAAAGCTTCTACAGAT	420	
Db	480	GAACTTTCTATCATCCAAAGTATGGGCTACAGAAACCGTGGCAAAAGCTTCTACAGAT	539	
OY	421	GAACCCGAAAATCCCTCCCTTCAGAGAAACAGCTCACTGTCACACTCTAACCTTGA	480	
Db	540	GAACCCGAAAATCCCTCCCTTCAGAGAAACAGCTCACTGTCACACTCTAACCTTGA	599	
OY	481	ACTGTGAGAACTGTGAGAGCAAAAGCAGCGGATACAACTCTCAAAAGAGCTGTCTACATT	540	
Db	600	ACTGTGAGAACTGTGAGAGCAAAAGCAGCGGATACAACTCTCAAAAGAGCTGTCTACATT	659	
OY	541	GAATTGGGATGTGATTTCTTCTGGAAGATACCGTTAATAAGGCACTTATTGCAAGTGTGGA	600	
Db	660	GAATTGGGATGTGATTTCTTCTGGAAGATACCGTTAATAAGGCACTTATTGCAAGTGTGGA	719	
OY	601	GATCAAGAAATGTTACAAATCACCCCTCAAGAAACAGAGATGAATAGTTGGATTCT	660	
Db	720	GATCAAGAAATGTTACAAATCACCCCTCAAGAAACAGAGATGAATAGTTGGATTCT	779	
OY	661	GCAAAAAAGGCTGCTGTGAAATTTTCTGAGACGAGATGAACAAATACTGAACATCATCAA	720	
Db	780	GCAAAAAAGGCTGCTGTGAAATTTTCTGAGACGAGATGAACAAATACTGAACATCATCAA	839	
OY	721	CCCAGTAATATGATTTTGAACACACAGAGCGTGCAGCTGAGAGCATCCAGAAAG	780	
Db	840	CCCAGTAATATGATTTTGAACACACAGAGCGTGCAGCTGAGAGCATCCAGAAAG	899	
OY	781	TATAGGAGTAGTGTGTTTCAAACTGTGATGTGAGAGCATGTGGACAAATATCATATGCC	840	
Db	900	TATAGGAGTAGTGTGTTTCAAACTGTGATGTGAGAGCATGTGGACAAATATCATATGCC	959	
OY	841	AGCTCATTTACAGATGAGAACAGCAGTATTATTACTACTAAAGACAGATGAATGTAGAA	900	
Db	960	AGCTCATTTACAGATGAGAACAGCAGTATTATTACTACTAAAGACAGATGAATGTAGAA	1019	
OY	901	AAGGCTGAATTCGTAAATPAAAAGCAAAAGCCTGCTTACGAGAGCCCAATPACAGA	960	
Db	1020	AAGGCTGAATTCGTAAATPAAAAGCAAAAGCCTGCTTACGAGAGCCCAATPACAGA	1079	
OY	961	TGGGCTGGAAGTAGGAACAATGTAATGATATGGGGGACTCCACACAGAAAAAAGCTA	1020	
Db	1080	TGGGCTGGAAGTAGGAACAATGTAATGATATGGGGGACTCCACACAGAAAAAAGCTA	1139	



OY	1021	GATCTGAAATGCGATACCCCTGGTGTAGAGAAAAAGATTAAGCAGAAACTGGCATGC	108
OY	1140	GATCTGAAATGCGATACCCCTGGTGTAGAGAAAAAGATTAAGCAGAAACTGGCATGC	1199
OY	1081	TCAGAGAAATCCAGAGACTACTCAAGATGTCCTCTGGATTAACCTAAATAGAGCAATTCAG	1140
Db	1200	TCAGAGAAATCCAGAGACTACTCAAGATGTCCTCTGGATTAACCTAAATAGAGCAATTCAG	1255
OY	1141	AAAGTTAATGAGTGGTTTTCCAGAGATGATGAATGTTAGTTCTGATGACTCACATGAT	1200
Db	1260	AAAGTTAATGAGTGGTTTTCCAGAGATGATGAATGTTAGTTCTGATGACTCACATGAT	1319
OY	1201	GGGGAGTCGTGATCAAAATGCCAAATAGCTGATGATTTGGAGCGTCTTAATAGAGGTAAAT	1260
Db	1320	GGGGAGTCGTGATCAAAATGCCAAATAGCTGATGATTTGGAGCGTCTTAATAGAGGTAAAT	1379
OY	1261	GAATATTCGTGTTCTTCAGAGAAATAGACTTACTGGCCAGTAGCTCCATAGAGCGTTTA	1320
Db	1380	GAATATTCGTGTTCTTCAGAGAAATAGACTTACTGGCCAGTAGCTCCATAGAGCGTTTA	1439
OY	1321	ATATGTAAAGTGAAAGAGTTCACCTCCAAATCAGTAGAGATTAATTTGAAAGACAAATA	1380
Db	1440	ATATGTAAAGTGAAAGAGTTCACCTCCAAATCAGTAGAGATTAATTTGAAAGACAAATA	1499
OY	1381	TTTGGGAAACCTATCCGAGAGAGGCAAGCCGCCCACTTAAGCCATGTAACTGAAAT	1440
Db	1500	TTTGGGAAACCTATCCGAGAGAGGCAAGCCGCCCACTTAAGCCATGTAACTGAAAT	1555
OY	1441	CTAATTTATAGAGCACTTTGTTACTGAGCCACAGATTAATACAAGAGCGTCCCTCCACAAT	1500
Db	1560	CTAATTTATAGAGCACTTTGTTACTGAGCCACAGATTAATACAAGAGCGTCCCTCCACAAT	1619
OY	1501	AAATTAAGCCGTTAAAGAGAGACCTTACATCAGGCGTTCATCCTGAGGATTTTATCAAGAA	1560
Db	1620	AAATTAAGCCGTTAAAGAGAGACCTTACATCAGGCGTTCATCCTGAGGATTTTATCAAGAA	1679
OY	1561	GCAGATTTGGGAGTTCAAAAGACACCTCCTAAATGATTAATCAAGGCACTAACCAAAACGAG	1620
Db	1680	GCAGATTTGGGAGTTCAAAAGACACCTCCTAAATGATTAATCAAGGCACTAACCAAAACGAG	1739
OY	1621	CAGAAATGCTCAAGTGAATATTAATCTAATATGCTGATGAGATTAATAAACAAGGTGAT	1680
Db	1740	CAGAAATGCTCAAGTGAATATTAATCTAATATGCTGATGAGATTAATAAACAAGGTGAT	1799
OY	1681	TCTATTCGAAATGAGAAAAATCCTTAACCCATAGATGATACCTCGAAAAAGATTCGCTTTC	1740
Db	1800	TCTATTCGAAATGAGAAAAATCCTTAACCCATAGATGATACCTCGAAAAAGATTCGCTTTC	1859
OY	1741	AAAAAGAAAGCTGAACCTTATTAAGCAGCAGTATTAAGCAATATGGAACCTGCAATTAATTC	1800
Db	1860	AAAAAGAAAGCTGAACCTTATTAAGCAGCAGTATTAAGCAATATGGAACCTGCAATTAATTC	1919
OY	1801	CACAATTTAAAGACACCTTAATAAAGAAATAGCTGAGAGAGAGTCTTACCAGGCATATTT	1860
Db	1920	CACAATTTAAAGACACCTTAATAAAGAAATAGCTGAGAGAGAGTCTTACCAGGCATATTT	1979
OY	1861	CATGGCGCTTGAACCTAGTAGTAGTAAGAAATCTAAGCCCACTTAATGTAGCAATTTGAA	1920
Db	1980	CATGGCGCTTGAACCTAGTAGTAGTAAGAAATCTAAGCCCACTTAATGTAGCAATTTGAA	2039
OY	1921	ATTGATATGTTGTTCTAGCAGTGAAGAGATTAAGAAAAAAAGTCAACCAATCCAGCTC	1980
Db	2040	ATTGATATGTTGTTCTAGCAGTGAAGAGATTAAGAAAAAAAGTCAACCAATCCAGCTC	2099
OY	1981	AGGCAACAGCAAAACCTTACAACCTCATGAGAGTTAAGACCTGCACACTGAGCCAAAG	2040
Db	2100	AGGCAACAGCAAAACCTTACAACCTCATGAGAGTTAAGACCTGCACACTGAGCCAAAG	2159
OY	2041	AGTAACAAGCCAAATGAACAGACAGTAAGTAAGACATGACATGATCTTCCAGAGCTG	2100
Db	2160	AGTAACAAGCCAAATGAACAGACAGTAAGTAAGACATGACATGATCTTCCAGAGCTG	2219

OY	2101	AAAGTAAACAATGCACCGCTGCTCTTTTCTAAAGTCTCAAAATPACAGTGAACCTTAAAGAA	2160
Db	2220	AAAGTTAAACAATGACCGCTGTTCTTTTCTAAAGTCTTCAAAATPACAGTGAACCTTAAAGAA	2279
OY	2161	TTTGTCAATCCTAGCCCTTCCAAAGAGAGAAAAAGAGAAACTAGAAACAGTTAAAGTG	2220
Db	2280	TTTGTCAATCCTAGCCCTTCCAAAGAGAGAAAAAGAGAAACTAGAAACAGTTAAAGTG	2339
OY	2221	TCTAATAATGCTGAAGACCCCAAGAATCTCAATGTAACGTGAGAAAGGGCTTTGCCAACT	2280
Db	2340	TCTAATAATGCTGAAGACCCCAAGAATCTCAATGTAACGTGAGAAAGGGCTTTGCCAACT	2399
OY	2281	GAAAGATCTGTAGAGAGTAGAGAGTATTTTCATTTGATGCTGGTACTGATATAGGACCTCAG	2340
Db	2400	GAAAGATCTGTAGAGAGTAGAGAGTATTTTCATTTGATGCTGGTACTGATATAGGACCTCAG	2459
OY	2341	GAAATATCTGCTTACTGGAAGTTAGCACTGTAGGAAAGGCAAAAACGAACCAATTA	2400
Db	2460	GAAATATCTGCTTACTGGAAGTTAGCACTGTAGGAAAGGCAAAAACGAACCAATTA	2519
OY	2401	TGTGTGAGTACGTGTGCAGCAATTTTGAACCCCAAGGAGCAATTCATGATGTTTCCAAA	2460
Db	2520	TGTGTGAGTACGTGTGCAGCAATTTTGAACCCCAAGGAGCAATTCATGATGTTTCCAAA	2579
OY	2461	GATAATAGCAATGAGACAGAAAGCCTTTAAGTATCCATTGGACATGAAAGTTAACCACT	2520
Db	2580	GATAATAGCAATGAGACAGAAAGCCTTTAAGTATCCATTGGACATGAAAGTTAACCACT	2639
OY	2521	CGGGAAACAGCATGTGAATATGGAAGTAAGTGAACCTGATGTCGATTTTCCAGATTA	2580
Db	2640	CGGGAAACAGCATGTGAATATGGAAGTAAGTGAACCTGATGTCGATTTTCCAGATTA	2699
OY	2581	TTTCAAGTTTCAAAGCGCCAGTCAATTGCTCTGTTTTCAAATCCAGAAATGACGAAGAG	2640
Db	2700	TTTCAAGTTTCAAAGCGCCAGTCAATTGCTCTGTTTTCAAATCCAGAAATGACGAAGAG	2759
OY	2641	GAATGTGCAACATTTCTCTGCCACTCTGCGGTCTTAAAGAAACAAGTCCAAAGTCACT	2700
Db	2760	GAATGTGCAACATTTCTCTGCCACTCTGCGGTCTTAAAGAAACAAGTCCAAAGTCACT	2819
OY	2701	TTTGAATGTGAACAAAAGAGAAATATCAAGAAAGATGATGCTAAATATCAACCTGTA	2760
Db	2820	TTTGAATGTGAACAAAAGAGAAATATCAAGAAAGATGATGCTAAATATCAACCTGTA	2879
OY	2761	CAGACAGTTAATATCTACTGACAGGCTTTCCTGTGTTGTCAGAAAGATTAAGCCAGTTGAT	2820
Db	2880	CAGACAGTTAATATCTACTGACAGGCTTTCCTGTGTTGTCAGAAAGATTAAGCCAGTTGAT	2939
OY	2821	AATGCGCAATGTAGTATCAAAAGAGGAGCCTGATGTTTTGTCTATCATCTCAGTTGAGAGGC	2880
Db	2940	AATGCGCAATGTAGTATCAAAAGAGGAGCCTGATGTTTTGTCTATCATCTCAGTTGAGAGGC	2999
OY	2881	AACGAAACTGAGCTCATTTACTCCAATAAACAATGAGCTTTTACAAAACCCATATCGTATA	2940
Db	3000	AACGAAACTGAGCTCATTTACTCCAATAAACAATGAGCTTTTACAAAACCCATATCGTATA	3059
OY	2941	CCACCACATTTTCCCATCACTCACTGTTTGTAAACTAAATCTAAGAAAAATCTGCTAGAG	3000
Db	3060	CCACCACATTTTCCCATCACTCACTGTTTGTAAACTAAATCTAAGAAAAATCTGCTAGAG	3119
OY	3001	GAAAACTTTGAGAAACATTCAAATCTCACAGCTGAAAGAGAAATGCGAAATGGAACATTC	3060
Db	3120	GAAAACTTTGAGAAACATTCAAATCTCACAGCTGAAAGAGAAATGCGAAATGGAACATTC	3179
OY	3061	AGTACAGTGAGACACAATTAGCCCGTAATTAACATTAGAGAAAAATGTTTTTAAAGAACCCAGC	3120
Db	3180	AGTACAGTGAGACACAATTAGCCCGTAATTAACATTAGAGAAAAATGTTTTTAAAGAACCCAGC	3239
OY	3121	TCAGCAATATTAATGAAGTGAAGTGTCCAGTACTAATGAAGGGCCCTCAGATTAATGA	3180
Db	3240	TCAGCAATATTAATGAAGTGAAGTGTCCAGTACTAATGAAGGGCCCTCAGATTAATGA	3299
OY	3181	ATAGGTTCCAGTGTGAAGAAACATTCACAGACGAGAACTAGTAGAAGAAACAAGGCCCAAAATTTG	3240



```
|||||
Db 3300 ATAGGTTCCAGTGTATATAAACAATTCAAGAGAACTGAGTGAAGAACAGAGGGCCAAATATG 3359
OY 3241 AATGATATGCTTAGATTAGGGGTTTTCGAACCTGAGGTCTATTAACAAAGTCTTCTGGG 3300
Db 3360 AATGCTATGCTTAGATTAGGGGTTTTCGAACCTGAGGTCTATTAACAAAGTCTTCTGGG 3419
OY 3301 AGTAATTTGAAGCATCTCTGAATATAAAGCAAGAAATATGAAGAACTAGTTCAGACTGT 3360
Db 3420 AGTAATTTGAAGCATCTCTGAATATAAAGCAAGAAATATGAAGAACTAGTTCAGACTGT 3479
OY 3361 AATACGATTTTCTCTCATATCTGATTTAGATTAACACGCTTATGGAGTACT 3420
Db 3480 AATACGATTTTCTCTCATATCTGATTTAGATTAACACGCTTATGGAGTACT 3539
OY 3421 CATGATCTCAGGTTTGTCTGAGACCTGATGACCTGTATGATGATGTTGAATTAAG 3480
Db 3540 CATGATCTCAGGTTTGTCTGAGACCTGATGACCTGTATGATGATGTTGAATTAAG 3599
OY 3481 GAAGATACATAGTTTCTGAAAAATGACATTAAAGAAAGTTCTGCTTTTTCAGCAAAAGC 3540
Db 3600 GAAGATACATAGTTTCTGAAAAATGACATTAAAGAAAGTTCTGCTTTTTCAGCAAAAGC 3659
OY 3541 GTCCAAAAAGAGAGCTTAGCAGAGAGTCTAGCCCTTTACCCATACACATTTGGCTCAG 3600
Db 3660 GTCCAAAAAGAGAGCTTAGCAGAGAGTCTAGCCCTTTACCCATACACATTTGGCTCAG 3719
OY 3601 GGTTCAGAGAGAGGGCCAAAGAAATTAGAGTCTCTGAAAGAGAACTTATCTGTGTGAGAT 3660
Db 3720 GGTTCAGAGAGAGGGCCAAAGAAATTAGAGTCTCTGAAAGAGAACTTATCTGTGTGAGAT 3779
OY 3661 GAAGAGCTTCCCTGCTTCCACACATTTGTTAATTTGTAAGTAAACAATATACCTTCTCAG 3720
Db 3780 GAAGAGCTTCCCTGCTTCCACACATTTGTTAATTTGTAAGTAAACAATATACCTTCTCAG 3839
OY 3721 TCTACTAGGCAATAGACCCGTGCTACCGAGTGTCTGTCTTAACAACACAGAGAAATTTA 3780
Db 3840 TCTACTAGGCAATAGACCCGTGCTACCGAGTGTCTGTCTTAACAACACAGAGAAATTTA 3899
OY 3781 TTATCTTGAAGAAATAGTTAAATGACTGCAATACAGTAAATATATGCAAAAGGATCT 3840
Db 3900 TTATCTTGAAGAAATAGTTAAATGACTGCAATACAGTAAATATATGCAAAAGGATCT 3959
OY 3841 CAGGAACATACCTTAGTAGAGAAACAAATGTTCTGTAAGTGTGTTTCTTACAGTGC 3900
Db 3960 CAGGAACATACCTTAGTAGAGAAACAAATGTTCTGTAAGTGTGTTTCTTACAGTGC 4019
OY 3901 ACTGAATTTGGAAGACTGTGACTGCAATATCAAAACACCCAGAGATCTTCTTGATGGTCT 3960
Db 4020 ACTGAATTTGGAAGACTGTGACTGCAATATCAAAACACCCAGAGATCTTCTTGATGGTCT 4079
OY 3961 TCCAAACAANTAGAGCATGCTGTAAGACCAAGGAGTGTGCTGAGTACAGAGAAATG 4020
Db 4080 TCCAAACAANTAGAGCATGCTGTAAGACCAAGGAGTGTGCTGAGTACAGAGAAATG 4139
OY 4021 GTTTCAGATGATGAAGAAAGAGAAAGGGCTTGGAAAGAAATATATCAAGAGCAAAAC 4080
Db 4140 GTTTCAGATGATGAAGAAAGAGAAAGGGCTTGGAAAGAAATATATCAAGAGCAAAAC 4199
OY 4081 ATGGAAATCAACTTATGCTGAACAGCATGTGGGTGTGAGAGTGAACAAGGCTCTCTGA 4140
Db 4200 ATGGAAATCAACTTATGCTGAACAGCATGTGGGTGTGAGAGTGAACAAGGCTCTCTGA 4259
OY 4141 GACTGCTCAGGGCTATCTCTCAGAGTGCATTTTAAACCACTCAGAGAGGATACCATG 4200
Db 4260 GACTGCTCAGGGCTATCTCTCAGAGTGCATTTTAAACCACTCAGAGAGGATACCATG 4319
OY 4201 CAACATTAACCTGATAAAGCTCCAGAGAAATGGCTGAAGCTGTGTAAGACAG 4260
Db 4320 CAACATTAACCTGATAAAGCTCCAGAGAAATGGCTGAAGCTGTGTAAGACAG 4379
OY 4261 CATGGAGCAGGCTTCTAAGCATACCTTTCATATAGTGAAGTCTTCTGCTTGAG 4320
|||||
Db 4380 CATGGAGCAGGCTTCTAAGCATACCTTTCATATAGTGAAGTACTTCTGCTTGAG 4439
OY 4321 GACCTCGAAATCCAGAACAAACCATCATGAGAAAGAGATTAATCTTCAGAAAGT 4380
Db 4440 GACCTCGAAATCCAGAACAAACCATCATGAGAAAGAGATTAATCTTCAGAAAGT 4499
OY 4381 AGTAATTAACCTATAAGCCAGAAATCCAGAGGCTTTCGCTGACAAAGTTTGAAGTGTCT 4440
Db 4500 AGTAATTAACCTATAAGCCAGAAATCCAGAGGCTTTCGCTGACAAAGTTTGAAGTGTCT 4559
OY 4441 GCAGATAGTTTCTACAGTAAATAAAGAACAGAGAGTGAAGAGTCAATCCCTCTTAA 4500
Db 4560 GCAGATAGTTTCTACAGTAAATAAAGAACAGAGAGTGAAGAGTCAATCCCTCTTAA 4619
OY 4501 TGCCCATCAATTTGATATAGTGTGATGAGTGTGATGAGTGTGATGAGTGTGATGAGTGTG 4560
Db 4620 TGCCCATCAATTTGATATAGTGTGATGAGTGTGATGAGTGTGATGAGTGTGATGAGTGTG 4679
OY 4561 AACTACCATCTCAGAGAGGCTCATTAAGTGTGATGAGTGTGATGAGTGTGATGAGTGTG 4620
Db 4680 AACTACCATCTCAGAGAGGCTCATTAAGTGTGATGAGTGTGATGAGTGTGATGAGTGTG 4739
OY 4621 GAGTGTGGCCACAGCATTTTGAAGCAACATCTTACTTCCAGAGCAAGCATCTAGAGGA 4680
Db 4740 GAGTGTGGCCACAGCATTTTGAAGCAACATCTTACTTCCAGAGCAAGCATCTAGAGGA 4799
OY 4681 ACCCCTTACTGGAATCTGGAATCAGCCCTTCTGATGAGTGTGATGAGTGTGATGAGTGTG 4740
Db 4800 ACCCCTTACTGGAATCTGGAATCAGCCCTTCTGATGAGTGTGATGAGTGTGATGAGTGTG 4859
OY 4741 GAAGACAGAGCCCCAGAGTCAAGTGTGTCACATCATCTTCAACCTCTGATG 4800
Db 4860 GAAGACAGAGCCCCAGAGTCAAGTGTGTCACATCATCTTCAACCTCTGATG 4919
OY 4801 AAAGTTCGCCAATTTGAAGTTGCAAGATCTGCCAGAGTGCAGTGTCTCTATCTACT 4860
Db 4920 AAAGTTCGCCAATTTGAAGTTGCAAGATCTGCCAGAGTGCAGTGTCTCTATCTACT 4979
OY 4861 GATACGTGCTGGGATTAATGATGAAGAAAGTGTGAGAGGAGAGGACAGAAATTTGACA 4920
Db 4980 GATACGTGCTGGGATTAATGATGAAGAAAGTGTGAGAGGAGAGGAGGAGGAGAAATTTGACA 5039
OY 4921 GCTTCAACAGAAAGGTTCAACAAAGAAATGTCATGTTGCTGTGCTGACCCAGAA 4980
Db 5040 GCTTCAACAGAAAGGTTCAACAAAGAAATGTCATGTTGCTGTGCTGACCCAGAA 5099
OY 4981 GAATTTATGCTGCTGTACAGATTTGCGCAAGAACACACATCATCTTAACTAATCTAAT 5040
Db 5100 GAATTTATGCTGCTGTACAGATTTGCGCAAGAACACACATCATCTTAACTAATCTAAT 5159
OY 5041 ACTGAAGAGTACTGATGATGTTATGAAGACAGATGCTGATGTTGTGTGAACGGACA 5100
Db 5160 ACTGAAGAGTACTGATGATGTTATGAAGACAGATGCTGATGTTGTGTGAACGGACA 5219
OY 5101 CTGAATATTTTCTAGGAATTCGGGAGAGAAATGGGTAGTATTTCTGGGTGACC 5160
Db 5220 CTGAATATTTTCTAGGAATTCGGGAGAGAAATGGGTAGTATTTCTGGGTGACC 5279
OY 5161 CAGTCTATTTAAAGAAAGAAATGCTGAATGACATGATTTTGAAGTCAAGAGAGATGTTG 5220
Db 5280 CAGTCTATTTAAAGAAAGAAATGCTGAATGACATGATTTTGAAGTCAAGAGAGATGTTG 5339
OY 5221 GTCAATGGAAGAACACCAAGGTCCAAAGCAGAGAAAGATTCAGAGACGAAAGATC 5280
Db 5340 GTCAATGGAAGAACACCAAGGTCCAAAGCAGAGAAAGATTCAGAGACGAAAGATC 5399
OY 5281 TTCAAGGGGCTTAAGAAATCTGTTGCTATGAGCCCTTACCAACATGCCACAGATCAAGTG 5340
Db 5400 TTCAAGGGGCTTAAGAAATCTGTTGCTATGAGCCCTTACCAACATGCCACAGATCAAGTG 5459
OY 5341 GAATGATGTGTACAGTGTGTGTCTGTGTGTGAAGAGCTTTCATCATTCACCTT 5400
Db 5460 GAATGATGTGTACAGTGTGTGTCTGTGTGTGAAGAGCTTTCATCATTCACCTT 5519
|||||
```

QY 5401 GGCACAGGTGTCCACCCCAATTTGGTGTGGAGCCAGATGCTTGACAGAGACATGSC 5460  
|||||  
Db 5520 GGCACAGGTGTCCACCCCAATTTGGTGTGGAGCCAGATGCTTGACAGAGACATGSC 5579  
QY 5461 TTCATGCAATTTGGGCAATGTGTGAGCACCCTGTGTGACCCGAGAGTGGTGTGGAC 5520  
|||||  
Db 5580 TTCATGCAATTTGGGCAATGTGTGAGCACCCTGTGTGACCCGAGAGTGGTGTGGAC 5639  
QY 5521 AGTGTAGCACTCTACCACTGTGAGAGAGCTGACACCTGCTGATACCCAGATCCCCAC 5580  
|||||  
Db 5640 AGTGTAGCACTCTACCACTGTGAGAGAGCTGACACCTGCTGATACCCAGATCCCCAC 5699  
QY 5581 AGCCACTAC 5589  
|||||  
Db 5700 AGCCACTAC 5708

RESULT 8  
ARI84044  
LOCUS ARI84044 5712 bp DNA Linear PAT 20-APR-2002  
DEFINITION Sequence 47 from patent US 6342483.  
ACCESSION ARI84044  
VERSION ARI84044.1 GI:20228013  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 5712)  
AUTHORS Holt,J.T., Jensen,R.A., Page,D.L., Obermiller,P.S.,  
Robinson-Benion,C.L. and Thompson,M.E.  
TITLE Method for detection and treatment of breast cancer  
JOURNAL Patent: US 6342483-A 47 29-JAN-2002;  
FEATURES  
source 1..5712  
location/Qualifiers  
BASE COUNT 1956 a 1099 c 1274 g 1383 t  
ORIGIN

Query Match 99.9%; Score 5585.8; DB 6; Length 5712;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 5587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGATTTATCTGCTCTTCGCGTTGAAGATACAAATGTCATTATGCTATGCAAAA 60  
|||||  
Db 120 ATGATTTATCTGCTCTTCGCGTTGAAGATACAAATGTCATTATGCTATGCAAAA 179  
QY 61 ATCTTAAGTGTCCCATCTGTCTGAGTGTATCAAGAACCTGTCTCCACAAAGTGTAC 120  
|||||  
Db 180 ATCTTAAGTGTCCCATCTGTCTGAGTGTATCAAGAACCTGTCTCCACAAAGTGTAC 239  
QY 121 CACATATTTTGCAAAATTTTGCACTGCTGAACTTCTCAACAGAAAGGCGCTTCACAG 180  
|||||  
Db 240 CACATATTTTGCAAAATTTTGCACTGCTGAACTTCTCAACAGAAAGGCGCTTCACAG 299  
QY 181 TGTCTTTATGTAAAGATGATATACCAAAAGAGCGCTTCAAGAAAGTACAGATTAGT 240  
|||||  
Db 300 TGTCTTTATGTAAAGATGATATACCAAAAGAGCGCTTCAAGAAAGTACAGATTAGT 359  
QY 241 CAACCTTTGAGAGCTATTGAAAATCATTTTGTCTTTTCACTTGTGACACAGGTTTGGAG 300  
|||||  
Db 360 CAACCTTTGAGAGCTATTGAAAATCATTTTGTCTTTTCACTTGTGACACAGGTTTGGAG 419  
QY 301 TATGCAAAAGCTATATTTTGGCAAAAAGGAAAATTAACCTCTGTAACATCTAAAGAT 360  
|||||  
Db 420 TATGCAAAAGCTATATTTTGGCAAAAAGGAAAATTAACCTCTGTAACATCTAAAGAT 479  
QY 361 GAAGTTTCTATCATCCAAAGATATGAGCTACAGAAACCGTGCACAAAGCTTCTACAGAGT 420  
|||||  
Db 480 GAAGTTTCTATCATCCAAAGATATGAGCTACAGAAACCGTGCACAAAGCTTCTACAGAGT 539  
QY 421 GAACCGCAAAATCTTCTCTGACAGAAACAGTCTCAGTGTCAACTCTTAACCTTGA 480  
|||||

Db 540 GAACCGCAAAATCTTCTCTGACAGAAACAGTCTCAGTGTCAACTCTTAACCTTGA 599  
QY 481 ACTGTGAGACCTTGAGGACAAAGCGGATTAACCTCAAAAAGAGCTGTCTCATTT 540  
|||||  
Db 600 ACTGTGAGACCTTGAGGACAAAGCGGATTAACCTCAAAAAGAGCTGTCTCATTT 659  
QY 541 GAATTTGGATCTGATTTCTTCTGAAGATACCGTTAATAAGCAACTTAATTCAGTGTGGG 600  
|||||  
Db 660 GAATTTGGATCTGATTTCTTCTGAAGATACCGTTAATAAGCAACTTAATTCAGTGTGGG 719  
QY 601 GATCAAGATTTGTTACAAATACCCCTCAGAGAACCGGATGAATCACTTTGATTTCT 660  
|||||  
Db 720 GATCAAGATTTGTTACAAATACCCCTCAGAGAACCGGATGAATCACTTTGATTTCT 779  
QY 661 GCAAAAAGGCTGCTGTGAATTTTCTGAGAGCGGATTAACCAATTAACATCAATCA 720  
|||||  
Db 780 GCAAAAAGGCTGCTGTGAATTTTCTGAGAGCGGATTAACCAATTAACATCAATCA 839  
QY 721 CCCAGTAATAATGATTTGAAACCCACAGAGAGCGTACAGTGAAGGATCCAGAAAG 780  
|||||  
Db 840 CCCAGTAATAATGATTTGAAACCCACAGAGAGCGTACAGTGAAGGATCCAGAAAG 899  
QY 781 TATCAGGAGTGTCTGTTTCAACTTCAGATGTGAGGACCATGTGACCAATTAATCAATGCC 840  
|||||  
Db 900 TATCAGGAGTGTCTGTTTCAACTTCAGATGTGAGGACCATGTGACCAATTAATCAATGCC 959  
QY 841 AGCTCATTTACAGCATGAGAACAGAGTTTATCTCACTAAACAGAAATGATAGAA 900  
|||||  
Db 960 AGCTCATTTACAGCATGAGAACAGAGTTTATCTCACTAAACAGAAATGATAGAA 1019  
QY 901 AAGGCTGAATTTCTGATTAATAAGCAACACCTGCGCTTACAGAGAGCCCAACATACAA 960  
|||||  
Db 1020 AAGGCTGAATTTCTGATTAATAAGCAACACCTGCGCTTACAGAGAGCCCAACATACAA 1079  
QY 961 TGGCTGTGAGTAAGAGAAACATGTAATGATAGCGGACCTCCAGCACAGAAAAAGGTA 1020  
|||||  
Db 1080 TGGCTGTGAGTAAGAGAAACATGTAATGATAGCGGACCTCCAGCACAGAAAAAGGTA 1139  
QY 1021 GATCTGAATGCTGATCCCTGTGTGAGAGAAAGAAATGAAATGAGAGAAATGCCATGC 1080  
|||||  
Db 1140 GATCTGAATGCTGATCCCTGTGTGAGAGAAAGAAATGAAATGAGAGAAATGCCATGC 1199  
QY 1081 TCAGAGAAATCTAGAGATCTGAAGATGTTCTTGGATAACACTTAATATAGACATTCAG 1140  
|||||  
Db 1200 TCAGAGAAATCTAGAGATCTGAAGATGTTCTTGGATAACACTTAATATAGACATTCAG 1259  
QY 1141 AAAGTTATGAGTGTGTTTCCAGAGAGATGAGTGTAGGTTCTGATGACTACATGAT 1200  
|||||  
Db 1260 AAAGTTATGAGTGTGTTTCCAGAGAGATGAGTGTAGGTTCTGATGACTACATGAT 1319  
QY 1201 GGGAGTCTGAATCAAAATGCCAAAGTGTGATGATTTGAGCTTCAATAGGATAGAT 1260  
|||||  
Db 1320 GGGAGTCTGAATCAAAATGCCAAAGTGTGATGATTTGAGCTTCAATAGGATAGAT 1379  
QY 1261 GAATATTTCTGTTCTTCAGAGAAATAGACTTACTGGCCAGTATCCTATGAGGCTTTA 1320  
|||||  
Db 1380 GAATATTTCTGTTCTTCAGAGAAATAGACTTACTGGCCAGTATCCTATGAGGCTTTA 1439  
QY 1321 ATATGTAAAGTGAAGAGTTCACCTCAAAATCACTGAGAGATATTTGAAGCAAAATA 1380  
|||||  
Db 1440 ATATGTAAAGTGAAGAGTTCACCTCAAAATCACTGAGAGATATTTGAAGCAAAATA 1499  
QY 1381 TTTGGGAAACCTATGGAAGAGGAGCGCTCCCACTTAAAGCATATGATCAATAAT 1440  
|||||  
Db 1500 TTTGGGAAACCTATGGAAGAGGAGCGCTCCCACTTAAAGCATATGATCAATAAT 1559  
QY 1441 CTATTTATGAGAGATTTTACTGAGCCACAGATATATCAAGAGGCTCCCTCAAAAT 1500  
|||||  
Db 1560 CTATTTATGAGAGATTTTACTGAGCCACAGATATATCAAGAGGCTCCCTCAAAAT 1619  
QY 1501 AAATTTAAGGCTAATAAGAGACCTACATCAGGCTTCATCTGAGGATTTATCAAGAA 1560  
|||||  
Db 1620 AAATTTAAGGCTAATAAGAGACCTACATCAGGCTTCATCTGAGGATTTATCAAGAA 1679

QY	1561	GCAGATTGCGAGTCTCAAAAGACTCTGAAATGATTAATCAGGCACTAAACCAACGGAG	1620
Db	1680	GCAGATTGCGAGTCTCAAAAGACTCTGAAATGATTAATCAGGCACTAAACCAACGGAG	1739
QY	1621	CAGAAATGTCAGTGAATGATATTACTAATATAGTGTGATGGAATTAACCAAAAGTGAT	1680
Db	1740	CAGAAATGTCAGTGAATGATATTACTAATATAGTGTGATGGAATTAACCAAAAGTGAT	1799
QY	1681	TCATTTCAGATGAGAAAAATCCTAACCCTAATAGATCAGTCCGAAAAAGATCGCTTC	1740
Db	1800	TCATTTCAGATGAGAAAAATCCTAACCCTAATAGATCAGTCCGAAAAAGATCGCTTC	1859
QY	1741	AAAAAGAAAGCTGAACCTATTAAGCAGAGATATAAGCAATATGAACTCGAATTTAAATATC	1800
Db	1860	AAAAAGAAAGCTGAACCTATTAAGCAGAGATATAAGCAATATGAACTCGAATTTAAATATC	1919
QY	1801	CACAATTCAAAAGCAGCTTAAAAAGATAGGCTGAGGAGGAAGTTCTTACGAGCATATT	1860
Db	1920	CACAATTCAAAAGCAGCTTAAAAAGATAGGCTGAGGAGGAAGTTCTTACGAGCATATT	1979
QY	1861	CATCGCTTGAACTAGTAGTACAGTGAAGAATCTAAGCCACCTAATTTACTGAATTCGA	1920
Db	1980	CATCGCTTGAACTAGTAGTACAGTGAAGAATCTAAGCCACCTAATTTACTGAATTCGA	2039
QY	1921	ATTGATAGTTGTTCTAGCAGTGAAGAGATAAAGAAAAAAAAGTACAACCAATGCCAGTC	1980
Db	2040	ATTGATAGTTGTTCTAGCAGTGAAGAGATAAAGAAAAAAAAGTACAACCAATGCCAGTC	2099
QY	1981	AGGCAACGCAAAAACCTTACAACTGATGAAGGTAAAGAACTGCACTGGAGCCAGAAG	2040
Db	2100	AGGCAACGCAAAAACCTTACAACTGATGAAGGTAAAGAACTGCACTGGAGCCAGAAG	2159
QY	2041	AGTACAAGCCAAATGAAACAGACAAAGTAAAAAGACATACAGTGTACTTTCCAGACTG	2100
Db	2160	AGTACAAGCCAAATGAAACAGACAAAGTAAAAAGACATACAGTGTACTTTCCAGACTG	2219
QY	2101	AAGTTAACCAATGACACCTGGTTCTTTACTAAGTGTCCAATACACAGTGAACCTTAAAGAA	2160
Db	2220	AAGTTAACCAATGACACCTGGTTCTTTACTAAGTGTCCAATACACAGTGAACCTTAAAGAA	2279
QY	2161	TTTGTCAATCTAGCCTTTCCAGAGAAAGAAAGAGAACTAGAAACAGTTAAAGTG	2220
Db	2280	TTTGTCAATCTAGCCTTTCCAGAGAAAGAAAGAGAACTAGAAACAGTTAAAGTG	2339
QY	2221	TCATATATGCTGGAAGACCCCAAAAGATCTCATGTTAGTGGAGAAAGGTTTTCCAACT	2280
Db	2340	TCATATATGCTGGAAGACCCCAAAAGATCTCATGTTAGTGGAGAAAGGTTTTCCAACT	2399
QY	2281	GAAAGATCTGTAGAGATAGCAGATTTCATTTGGTACTGGTACTGATTTAGGCATCAG	2340
Db	2400	GAAAGATCTGTAGAGATAGCAGATTTCATTTGGTACTGGTACTGATTTAGGCATCAG	2459
Y	2341	GAAAGTATCTCGTTACTGGAAGTTAGCAGCTGTAGGGAGGCAAAAACAGAACCAATATA	2400
Db	2460	GAAAGTATCTCGTTACTGGAAGTTAGCAGCTGTAGGGAGGCAAAAACAGAACCAATATA	2519
QY	2401	TGATGTAGTGAAGTGTGAGCATTTGAAACCCCAAGGACATTAATTCATGGTTGTTCCAA	2460
Db	2520	TGATGTAGTGAAGTGTGAGCATTTGAAACCCCAAGGACATTAATTCATGGTTGTTCCAA	2579
QY	2461	GATTAATGAAATGACACAGAGGCTTTAAGTATCCATTGGACATGGAAGTTAACACAGT	2520
Db	2580	GATTAATGAAATGACACAGAGGCTTTAAGTATCCATTGGACATGGAAGTTAACACAGT	2639
QY	2521	CGGGAACAAGCATATAGAAATGGAAGAAAGTGAACCTTGATGCTCATGTTTCGACATATCA	2580
Db	2640	CGGGAACAAGCATATAGAAATGGAAGAAAGTGAACCTTGATGCTCATGTTTCGACATATCA	2699
QY	2581	TTTCAAGGTTTCAAAAGCCGACGTCAATTTGCTGTGTTTTCAAATCCAGAAATGCAGAAAG	2640
Db	2700	TTTCAAGGTTTCAAAAGCCGACGTCAATTTGCTGTGTTTTCAAATCCAGAAATGCAGAAAG	2759

QY	2641	GAATGTGCAACATTTCTCTGCCACTCTGTGGTCTTTAAAGAAACAAGTCCAAAGTCACT	2700
Db	2760	GAATGTGCAACATTTCTCTGCCACTCTGTGGTCTTTAAAGAAACAAGTCCAAAGTCACT	2819
QY	2701	TTTGAATGTGAACAAAGGAAGAAATCAAGAAAGATGTCCTAATATCAACCTGTGA	2760
Db	2820	TTTGAATGTGAACAAAGGAAGAAATCAAGAAAGATGTCCTAATATCAACCTGTGA	2879
QY	2761	CAGACAGTTAATATCATCTGACAGGCTTTCCTGTGGTGGTGCAGAAAGATATACCCAGTTGAT	2820
Db	2880	CAGACAGTTAATATCATCTGACAGGCTTTCCTGTGGTGGTGCAGAAAGATATACCCAGTTGAT	2939
QY	2821	AATGCCAATGTAGTATCAACAGAGGCTCTAGTTTTGTCTATCATCTCACTTCAGAGGC	2880
Db	2940	AATGCCAATGTAGTATCAACAGAGGCTCTAGTTTTGTCTATCATCTCACTTCAGAGGC	2999
QY	2881	AACGAAATGTGAGTCATCTACTCCAAATPAACATGAGCTTTTACAAAACCCATATCGTATA	2940
Db	3000	AACGAAATGTGAGTCATCTACTCCAAATPAACATGAGCTTTTACAAAACCCATATCGTATA	3059
QY	2941	CCACCACCTTTTCCCATCAAGTCACTTTTGTAAAACTAAATGTAGAAAAATCTGCTAGAG	3000
Db	3060	CCACCACCTTTTCCCATCAAGTCACTTTTGTAAAACTAAATGTAGAAAAATCTGCTAGAG	3119
QY	3001	GAAACCTTTGAGGAACATTCATATGTCACTGTAAGAGAAATGGAAATGAGACATTTCCA	3060
Db	3120	GAAACCTTTGAGGAACATTCATATGTCACTGTAAGAGAAATGGAAATGAGACATTTCCA	3179
QY	3061	AGTCAGTGCACCAATATGAGCCGTAATPACCTTAGAGAAATGTTTTAAAGAACCCAGC	3120
Db	3180	AGTCAGTGCACCAATATGAGCCGTAATPACCTTAGAGAAATGTTTTAAAGAACCCAGC	3239
QY	3121	TCAAGCAATATTAATGAAGTAGTGTCCAGTACTATGTAAGTGGGCTCCAGTATTAATGAA	3180
Db	3240	TCAAGCAATATTAATGAAGTAGTGTCCAGTACTATGTAAGTGGGCTCCAGTATTAATGAA	3299
QY	3181	ATAGGTTCCAGTGATGAAACATTCACAGACACTAGGTAGAAACAGAGGCCAAATATGG	3240
Db	3300	ATAGGTTCCAGTGATGAAACATTCACAGACACTAGGTAGAAACAGAGGCCAAATATGG	3359
QY	3241	AATGCTATGCTTACATTTAGGGGTTTTGCAACCTGAGGCTTAATACAAAGCTTCCTCGGA	3300
Db	3360	AATGCTATGCTTACATTTAGGGGTTTTGCAACCTGAGGCTTAATACAAAGCTTCCTCGGA	3419
QY	3301	AGTAATTTGTAGCAATCTCCGAAATTAAGAAAGCAAAATATGAGAAGTAGTTCAACACTGT	3360
Db	3420	AGTAATTTGTAGCAATCTCCGAAATTAAGAAAGCAAAATATGAGAAGTAGTTCAACACTGT	3479
QY	3361	AATACAGATTTCTCTCCATATCTGATTTAGATTAAGTATGACAGCCCTATGGGAAGTAGT	3420
Db	3480	AATACAGATTTCTCTCCATATCTGATTTAGATTAAGTATGACAGCCCTATGGGAAGTAGT	3539
QY	3421	CATGATCTCAGGTTTTGTTCTGAGACACCTGATGACCTGTTAGATGATGGTGAATAAG	3480
Db	3540	CATGATCTCAGGTTTTGTTCTGAGACACCTGATGACCTGTTAGATGATGGTGAATAAG	3599
QY	3481	GAAGATCTAGATTTTGTGTAAGAAATGACATTAAGGAAATTTCTGCTGTTTTAGCAAAAGC	3540
Db	3600	GAAGATCTAGATTTTGTGTAAGAAATGACATTAAGGAAATTTCTGCTGTTTTAGCAAAAGC	3659
QY	3541	GTCCAGAAAGAGAGCTTAGCAGGAGTCTTAGCCCTTTACCCCATATACATTTTGGCTCAG	3600
Db	3660	GTCCAGAAAGAGAGCTTAGCAGGAGTCTTAGCCCTTTACCCCATATACATTTTGGCTCAG	3719
QY	3601	GGTTACGAGAGAGGGGCCAAGAAATTAGAGTCTCAGAGAGAACTTATCTAGTGAGGAT	3660
Db	3720	GGTTACGAGAGAGGGGCCAAGAAATTAGAGTCTCAGAGAGAACTTATCTAGTGAGGAT	3779
QY	3661	GAAAGGCTTCCCTGCTTCCACACTTGTATTGTTGGTAAAGTAACATPATACCTTCAG	3720
Db	3780	GAAAGGCTTCCCTGCTTCCACACTTGTATTGTTGGTAAAGTAACATPATACCTTCAG	3839
QY	3721	TCTACTAGGCACTTACACCGTGTCTACGAGTGTCTGCTATAAGAACACAGAGGAATTTTA	3780

|||||  
Db 3840 TCTACTAGCATAGCACCCTGGTCTACCGAGTGTCTGTAAAGAACAGAGGAGAAATTTA 3899  
OY 3781 TTATCATTTGAAGAAATAGCTTAAATGCTGACGATACACAGTAATATTTGGCAAGGCATCT 3840  
Db 3900 TTATCATTTGAAGAAATAGCTTAAATGCTGACGATACACAGTAATATTTGGCAAGGCATCT 3959  
OY 3841 CAGGAACATCAGCTTAGTGAAGAAACAAATGTTGTGCTAGCTGTGTTTCTTCCACAGTGC 3900  
Db 3960 CAGGAACATCAGCTTAGTGAAGAAACAAATGTTGTGCTAGCTGTGTTTCTTCCACAGTGC 4019  
OY 3901 AGTGAATTTGAAGAACTTGACGCAAAATACAAACACCAGAGATCCTTTCTGATTTGCTCT 3960  
Db 4020 AGTGAATTTGAAGAACTTGACGCAAAATACAAACACCAGAGATCCTTTCTGATTTGCTCT 4079  
OY 3961 TCCAAACAATGAGCATGCTGTAAGCAAGGAGGTTGCTGTGCTGACAGAAATTTG 4020  
Db 4080 TCCAAACAATGAGCATGCTGTAAGCAAGGAGGTTGCTGTGCTGACAGAAATTTG 4139  
OY 4021 GTTTCAGATGATGAAGAAAGAGAAACGGCTTGAAGAAATATATCAAGAGCAAAAGC 4080  
Db 4140 GTTTCAGATGATGAAGAAAGAGAAACGGCTTGAAGAAATATATCAAGAGCAAAAGC 4199  
OY 4081 ATGATTTCAAACTTAGGTGAGCAGCATCTGGGTGTGAGAGTGAACAAGCGTCTGTGA 4140  
Db 4200 ATGATTTCAAACTTAGGTGAGCAGCATCTGGGTGTGAGAGTGAACAAGCGTCTGTGA 4259  
OY 4141 GACTCTCAGGAGCTTTCCTCTGAGAGTGATTTTAAACCATCAGAGAGGATATACCATG 4200  
Db 4260 GACTCTCAGGAGCTTTCCTCTGAGAGTGATTTTAAACCATCAGAGAGGATATACCATG 4319  
OY 4201 CAACATAACTGTATAAGCTCCAGCAGAAATGGCTGAACCTAGAACTGTGTAAGACAG 4260  
Db 4320 CAACATAACTGTATAAGCTCCAGCAGAAATGGCTGAACCTAGAACTGTGTAAGACAG 4379  
OY 4261 CATGGAGCCAGCCTTTCTAACAGCTACCTTCCATCATTAAGTACTCTTCTGCTTGGAG 4320  
Db 4380 CATGGAGCCAGCCTTTCTAACAGCTACCTTCCATCATTAAGTACTCTTCTGCTTGGAG 4439  
OY 4321 GACCTGCGCAATTCAGAAACAAGCATCAGAAAAAGCATTTAACTTACAGAAAAAT 4380  
Db 4440 GACCTGCGCAATTCAGAAACAAGCATCAGAAAAAGCATTTAACTTACAGAAAAAT 4499  
OY 4381 AGTGAATACCTTATAAGCCAGAAATCCAGAAAGCCTTTCTGTACAAAGTTTGAGGTCT 4440  
Db 4500 AGTGAATACCTTATAAGCCAGAAATCCAGAAAGCCTTTCTGTACAAAGTTTGAGGTCT 4559  
OY 4441 GCAGTACTTCTACAGTAAATAATAAGAACAGAGTGAAGAGTCAATCCCTCTTAA 4500  
Db 4560 GCAGTACTTCTACAGTAAATAATAAGAACAGAGTGAAGAGTCAATCCCTCTTAA 4619  
OY 4501 TGCCCATATTAGATGATAGGTGTACATGACAGTGTGCTGGAGTCTTCAAGATAGA 4560  
Db 4620 TGCCCATATTAGATGATAGGTGTACATGACAGTGTGCTGGAGTCTTCAAGATAGA 4679  
OY 4561 AACTRACCATCTCAAGAGAGCTCATTAAGTTGTTGATGTGAGAGCAACAGCTGAA 4620  
Db 4680 AACTRACCATCTCAAGAGAGCTCATTAAGTTGTTGATGTGAGAGCAACAGCTGAA 4739  
OY 4621 GAGTCTGGGCCACAGATTTGACGGAACATCTTACTTCCCAAGGCAACATTAAGGGGA 4680  
Db 4740 GAGTCTGGGCCACAGATTTGACGGAACATCTTACTTCCCAAGGCAACATTAAGGGGA 4799  
OY 4681 ACCCTTACCTGGAATCTGAATCAGCTCTTCTGTGATGACCTGGAATCTGATCTTCT 4740  
Db 4800 ACCCTTACCTGGAATCTGAATCAGCTCTTCTGTGATGACCTGGAATCTGATCTTCT 4859  
OY 4741 GAAGACAGAGCCCGAGAGTCACTGTGTGCAACATACATCTTCAACCTCTGCAATG 4800  
Db 4860 GAAGACAGAGCCCGAGAGTCACTGTGTGCAACATACATCTTCAACCTCTGCAATG 4919  
OY 4801 AAAGTCCCAATTGAAGTTGCAAAATCTGCCAGAGTCCAGCTGCTCATACTACT 4860  
|||||

Db 4920 AAAGTCCCAATTGAAGTTGCAAAATCTGCCAGAGTCCAGCTGCTCATACTACT 4979  
OY 4861 GATACCTGCTGGGTATATATCATTTGAAGAAATGTGACGAGGAGAAACCCAGAAATTTGCA 4920  
Db 4980 GATACCTGCTGGGTATATATCATTTGAAGAAATGTGACGAGGAGAAACCCAGAAATTTGCA 5039  
OY 4921 GCTTCAACAGAAAGGCTCAACAAAGAAATGTCCATGCTGTGTGCTGAGCCCTGACCCAGAA 4980  
Db 5040 GCTTCAACAGAAAGGCTCAACAAAGAAATGTCCATGCTGTGTGCTGAGCCCTGACCCAGAA 5099  
OY 4981 GAATTTATCTGCTGTACAGATTTGCCAGAAACACCATCATCTTAACTTAATTAAT 5040  
Db 5100 GAATTTATCTGCTGTACAGATTTGCCAGAAACACCATCATCTTAACTTAATTAAT 5159  
OY 5041 ACTGAAGACATCACTATGCTGTATGAAGAAAGATGCTGAGTGTGCTGAGAGGCA 5100  
Db 5160 ACTGAAGACATCACTATGCTGTATGAAGAAAGATGCTGAGTGTGCTGAGAGGCA 5219  
OY 5101 CTGAATATTTCTGTAGAAATTCGGGAGAGAAATGGTAGTTAGCTATTCTGGGTGACC 5160  
Db 5220 CTGAATATTTCTGTAGAAATTCGGGAGAGAAATGGTAGTTAGCTATTCTGGGTGACC 5279  
OY 5161 CAGCTTATTTAAAGAAAGAAATGCTGAATGACATGATTTGAAGTCAAGAGATGTG 5220  
Db 5280 CAGCTTATTTAAAGAAAGAAATGCTGAATGACATGATTTGAAGTCAAGAGATGTG 5339  
OY 5221 GTCAATGAGAAAGAACCCAGGATGCTCAAGAGGAGCAAGAAATCCAGAGAAAGATC 5280  
Db 5340 GTCAATGAGAAAGAACCCAGGATGCTCAAGAGGAGCAAGAAATCCAGAGAAAGATC 5399  
OY 5281 TTCAGGGGCTAGAAATCTGTTGCTATGAGGCCCTTACCAACATGCCCCAGATCAACTG 5340  
Db 5400 TTCAGGGGCTAGAAATCTGTTGCTATGAGGCCCTTACCAACATGCCCCAGATCAACTG 5459  
OY 5341 GAATGATGTGACAGCTGTGTGCTGTCTGTGTGGAAGAGCTTTATCATTTACCTT 5400  
Db 5460 GAATGATGTGACAGCTGTGTGCTGTGTGGAAGAGCTTTATCATTTACCTT 5519  
OY 5401 GGCACAGGTGTCACCAATTTGTTGTTGTCAGCAGATGCTGTGACAGAGAAATGGC 5460  
Db 5520 GGCACAGGTGTCACCAATTTGTTGTTGTCAGCAGATGCTGTGACAGAGAAATGGC 5579  
OY 5461 TTCATGCAATTTGGCAGATGTGTGAGGACCTGTGTGACCCGAGAGTGGGTGGAC 5520  
Db 5580 TTCATGCAATTTGGCAGATGTGTGAGGACCTGTGTGACCCGAGAGTGGGTGGAC 5639  
OY 5521 AGTGAAGCCTCTACAGAGTCCAGAGAGCTGAGACACTTACCTGATACCCCAATCCCCAC 5580  
Db 5640 AGTGAAGCCTCTACAGAGTCCAGAGAGCTGAGACACTTACCTGATACCCCAATCCCCAC 5699  
OY 5581 AGCCACTAC 5589  
Db 5700 AGCCACTAC 5708  
  
RESULT 9  
AR004673  
LOCUS AR004673 5914 bp DNA Linear PAT 04-DEC-1998  
DEFINITION Sequence 1 from patent US 5747282.  
ACCESSION AR004673  
VERSION AR004673.1 GI:3965552  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 5914)  
AUTHORS Skolnick,M.H., Goldgar,D.E., Miki,Y., Swenson,J., Kamb,A.,  
Hershenov,K.D., Shattuck-Eidens,D.M., Tavtigian,S.V., Wiseman,R.W.  
and Futreal,P.Andrew.  
170-linked breast and ovarian cancer susceptibility gene  
JOURNAL Patent: US 5747282-A 1 05-MAY-1998;  
FEATURES  
Location/Qualifiers  
source 1..5914

BASE COUNT 2006 a 1156 c 1316 g 1436 t  
ORIGIN /organism="unknown"  
Query Match 99.9%; Score 5585.8; DB 6; Length 5914;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 5387; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 ATGATTTATCTGCTCTGCGTTGGAAGATCAAAATGCTATATGCTATGACGAAA 60  
120 ATGATTTATCTGCTCTGCGTTGGAAGATCAAAATGCTATATGCTATGACGAAA 179  
DB 61 ATCTAGAGTGTCCATCTGCTGAGTTGATCAAGAACTGTCTCCAAAGTGTGAC 120  
180 ATCTAGAGTGTCCATCTGCTGAGTTGATCAAGAACTGTCTCCAAAGTGTGAC 229  
QY 121 CACATATTTTGCAAAATTTTGCAATGCTGAAACTTCTACCCAGAGAAAGGCGCTTACAG 180  
240 CACATATTTTGCAAAATTTTGCAATGCTGAAACTTCTACCCAGAGAAAGGCGCTTACAG 299  
DB 181 TGTCTTTATGTAGATGATATATACCAAAAGAGGCTACAAAGATGAGATTTAGT 240  
300 TGTCTTTATGTAGATGATATATACCAAAAGAGGCTACAAAGATGAGATTTAGT 359  
QY 241 CAACCTGTGGAAGAGCTATTTGAAATCATTTGTCTTTTCAAGCTTGACACAGGTTGGAG 300  
360 CAACCTGTGGAAGAGCTATTTGAAATCATTTGTCTTTTCAAGCTTGACACAGGTTGGAG 419  
DB 301 TATGCAACAGCTATATTTTGGCAAAAGAAATTAACCTCTCTGACATCTTAAAGAT 360  
420 TATGCAACAGCTATATTTTGGCAAAAGAAATTAACCTCTCTGACATCTTAAAGAT 479  
QY 361 GAAGTTTCTATCATCCAAAGTATGAGCTCAGAAACGCTGCCAAAGATCTTACAGAGT 420  
480 GAAGTTTCTATCATCCAAAGTATGAGCTCAGAAACGCTGCCAAAGATCTTACAGAGT 539  
DB 421 GAACCCGAAATCTCTTCTGCGAGAAACAGCTCTGATGCTCAACTCTTAACTTGA 480  
540 GAACCCGAAATCTCTTCTGCGAGAAACAGCTCTGATGCTCAACTCTTAACTTGA 599  
QY 481 ACTGTGAGAACTCTGAGCAACAAAGCGGATATCAACCTCAAAAGAGCTGTCTACAT 540  
600 ACTGTGAGAACTCTGAGCAACAAAGCGGATATCAACCTCAAAAGAGCTGTCTACAT 659  
DB 541 GAATGGAGATCTGATCTCTGAGATACCGTTAATAGCAACTATTTGCAAGTGGGA 600  
660 GAATGGAGATCTGATCTCTGAGATACCGTTAATAGCAACTATTTGCAAGTGGGA 719  
QY 601 GATCAAGATTTGTACAAATCAACCCCTCAAGGAACCGAGGATGAATTCAGTTGATCT 660  
720 GATCAAGATTTGTACAAATCAACCCCTCAAGGAACCGAGGATGAATTCAGTTGATCT 779  
DB 661 GCAAAAAAGGCTGCTTGTGAATTTTCTGAGAGGATGTACAAATATCTGACATCATCA 720  
780 GCAAAAAAGGCTGCTTGTGAATTTTCTGAGAGGATGTACAAATATCTGACATCATCA 839  
QY 721 CCCAGTAATATGATTTGAAACACACTGAGAAAGCGTGCAGAGGAGATCCAGAAAG 780  
840 CCCAGTAATATGATTTGAAACACACTGAGAAAGCGTGCAGAGGAGATCCAGAAAG 899  
DB 841 TATCAGGGAGTCTGTTTCAAACTTGCATGTGAGCCATGTGGCAAAATCTCATGCC 959  
900 TATCAGGGAGTCTGTTTCAAACTTGCATGTGAGCCATGTGGCAAAATCTCATGCC 959  
QY 901 AACCTATTACAGATGAGAAACAGAGTTTATATCTACAAAGAGATGAATGTAGAA 900  
960 AACCTATTACAGATGAGAAACAGAGTTTATATCTACAAAGAGATGAATGTAGAA 1019  
DB 901 AAGGCTGAATTTGTATATAAAGCAAAACAGGCTGGTTAGCAAGAGCCAACTATAACAA 960  
1020 AAGGCTGAATTTGTATATAAAGCAAAACAGGCTGGTTAGCAAGAGCCAACTATAACAA 1079  
QY 961 TGGGCTGGAAGTAAAGAAACATGTAATGATAGCGGAGCTCCAGACAGAAAAAAGGTA 1020

DB 1080 TGGGCTGGAAGTAAAGAAACATGTAATGATAGCGGAGCTCCAGACAGAAAAAAGGTA 1139  
QY 1021 GATCTAAATGCTGATCCCTGCTGAGAGAAAGAAATGATATAGCAAGAAACGCTATGC 1080  
DB 1140 GATCTAAATGCTGATCCCTGCTGAGAGAAAGAAATGATATAGCAAGAAACGCTATGC 1199  
QY 1081 TCAGAAATCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140  
1200 TCAGAAATCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1259  
DB 1141 AAAGTTAATGAGTGTGTTTCCAGAAATGATGATGATGATGATGATGATGATGATGAT 1200  
1260 AAAGTTAATGAGTGTGTTTCCAGAAATGATGATGATGATGATGATGATGATGATGAT 1319  
QY 1201 GGGGAGTCTGATCAATATGCTGCAAGAGTATGATGATGATGATGATGATGATGATGAT 1260  
1320 GGGGAGTCTGATCAATATGCTGCAAGAGTATGATGATGATGATGATGATGATGATGAT 1379  
DB 1261 GAATATTCTGCTTCTCAGAGAAATGAGACTTCTGCGCAGATGATGATGATGATGATGAT 1320  
1380 GAATATTCTGCTTCTCAGAGAAATGAGACTTCTGCGCAGATGATGATGATGATGATGAT 1439  
QY 1321 ATATGTAAAGTGAAGAGTCTCACTCAAAATCAGTAGAGATATATTTGAAGCAAAATA 1380  
1440 ATATGTAAAGTGAAGAGTCTCACTCAAAATCAGTAGAGATATATTTGAAGCAAAATA 1499  
DB 1381 TTTGGGAAACCTATGCGAAGAAAGCAAGGCTCCCACTTAAAGCATGATGATGAT 1440  
1500 TTTGGGAAACCTATGCGAAGAAAGCAAGGCTCCCACTTAAAGCATGATGATGATGAT 1559  
QY 1441 CTATATTATGAGACATTTGTTTACTGAGCCACAGATATATCAAGAGGCTCCCAAAAT 1500  
1560 CTATATTATGAGACATTTGTTTACTGAGCCACAGATATATCAAGAGGCTCCCAAAAT 1619  
DB 1501 AAATTAAGCGTAAAGAGAGACCTACATCAGGCTTCTATCTGAGATTTTATCAAGAAA 1560  
1620 AAATTAAGCGTAAAGAGAGACCTACATCAGGCTTCTATCTGAGATTTTATCAAGAAA 1679  
QY 1561 GCAGATTTGGCACTTAAAAAGACTCCTGAAATGATATATCAGGAGACTAACCAAGGAG 1620  
1680 GCAGATTTGGCACTTAAAAAGACTCCTGAAATGATATATCAGGAGACTAACCAAGGAG 1739  
DB 1621 CAGAAATGCTCAGTGTGATGATATATTAATAGTGTGATGAGATTAACCAAGAGTGT 1680  
1740 CAGAAATGCTCAGTGTGATGATATATTAATAGTGTGATGAGATTAACCAAGAGTGT 1799  
QY 1681 TCTATTCAAGATGAGAAAAATCTTAACCAATGAAATCACTGAAAAAGATCTGCTTTC 1740  
1800 TCTATTCAAGATGAGAAAAATCTTAACCAATGAAATCACTGAAAAAGATCTGCTTTC 1859  
DB 1741 AAAAGCAAAAGCTGACCTATTAAGCAGCAGATATAGCAATATGAGACTCGAATTAATATC 1800  
1860 AAAAGCAAAAGCTGACCTATTAAGCAGCAGATATAGCAATATGAGACTCGAATTAATATC 1919  
QY 1801 CACAATTCAAAAGCACCCTAAAGAAATAGGCTGAGAGGAGTCTTACAGAGCATATT 1860  
1920 CACAATTCAAAAGCACCCTAAAGAAATAGGCTGAGAGGAGTCTTACAGAGCATATT 1979-  
DB 1861 CATGCGCTTGAACCTAGTACGTAGATAAATCTTAAGCCCACTAATTTGTAATTCGAA 1920  
1980 CATGCGCTTGAACCTAGTACGTAGATAAATCTTAAGCCCACTAATTTGTAATTCGAA 2039  
QY 1921 ATTTGATAGTGTCTTCTAGCAGTGAAGAGATTAAGAAAAAAGTACCAACCAATGCGAGTC 1980  
2040 ATTTGATAGTGTCTTCTAGCAGTGAAGAGATTAAGAAAAAAGTACCAACCAATGCGAGTC 2099  
DB 1981 AGGCACAGCAGAAACCTTACAGCTCATGGAAGGTAAAGCACTGCAACTGAGCAGAGAG 2040  
2100 AGGCACAGCAGAAACCTTACAGCTCATGGAAGGTAAAGCACTGCAACTGAGCAGAGAG 2159  
QY 2041 AGTAAACCAAGCAATGAAACAGCAAGTAAAGACATGACAGTATCTTCCAGAGACTG 2100

Dp	2160	AGTAAACAAGCCAAATGAAACACAGCAAGTAAAGACATGACACGGATCTTTCCAGAGCTG	2219
Oy	2101	AAGTTAAACAATGACACCTGGTTCTTTTACTAAGTGTTCAAATPACCAGTGAACCTTAAAGAA	2160
Dp	2220	AAGTTAAACAATGACACCTGGTTCTTTTACTAAGTGTTCAAATPACCAGTGAACCTTAAAGAA	2279
Oy	2161	TTTGTCAATCTTAGCCCTTCCAAAGAAACAAAAAAGAAACACTTGAACAGGTTAAAGTG	2220
Dp	2280	TTTGTCAATCTTAGCCCTTCCAAAGAAACAAAAAAGAAACACTTGAACAGGTTAAAGTG	2339
Oy	2221	TCTAATTAATGCTGAAGACCCCAAGATGCTCATGTGTTAAATGGAGAAAGGTTTCCAACT	2280
Dp	2340	TCTAATTAATGCTGAAGACCCCAAGATGCTCATGTGTTAAATGGAGAAAGGTTTCCAACT	2399
Oy	2291	GAAGATCTGTAGAGATAGCAGATATTTTCATTGTGTAACCTGGTACTGATTAATGGCACTCAG	2340
Dp	2400	GAAGATCTGTAGAGATAGCAGATATTTTCATTGTGTAACCTGGTACTGATTAATGGCACTCAG	2459
Oy	2341	GAAGATATCTGTTTACTGTGAAGTTTATGCACTATAGGGAAGCCAAAACAGAACCAATPAA	2400
Dp	2460	GAAGATATCTGTTTACTGTGAAGTTTATGCACTATAGGGAAGCCAAAACAGAACCAATPAA	2519
Oy	2401	TGTGTGATGCTAGTGTGCAGCAATTTGAAAAACCCCAAGGACATTAATTCATGGTTGTCCAA	2460
Dp	2520	TGTGTGATGCTAGTGTGCAGCAATTTGAAAAACCCCAAGGACATTAATTCATGGTTGTCCAA	2579
Oy	2461	GATATATAGAAATGACACAGAAAGGCTTTAAGTATCCATTGGGACATGAAAGTTAACACAGT	2520
Dp	2580	GATATATAGAAATGACACAGAAAGGCTTTAAGTATCCATTGGGACATGAAAGTTAACACAGT	2639
Oy	2521	CGGGAACAACATATAGAAATGGAAAGAAAGTAACTTGATGCTCAGTATTTGACAAATPAA	2580
Dp	2640	CGGGAACAACATATAGAAATGGAAAGAAAGTAACTTGATGCTCAGTATTTGACAAATPAA	2699
Oy	2581	TTTCAAGGTTTCAAAAGCCGCACTCATATTTGCTGTTTTCAATPCCAGAAATGACAGAAAG	2640
Dp	2700	TTTCAAGGTTTCAAAAGCCGCACTCATATTTGCTGTTTTCAATPCCAGAAATGACAGAAAG	2759
Oy	2641	GAATGTGCAACATTTCTGTGCCCTCTGTGGTCTTTAAAGAAACAAGTCCAAAAGTCACT	2700
Dp	2760	GAATGTGCAACATTTCTGTGCCCTCTGTGGTCTTTAAAGAAACAAGTCCAAAAGTCACT	2819
Oy	2701	TTTGAATGTGAACAAAGGAAGAAATATCAAGAAAGATGAGTAAATATCAAGCCCTGTA	2760
Dp	2820	TTTGAATGTGAACAAAGGAAGAAATATCAAGAAAGATGAGTAAATATCAAGCCCTGTA	2879
Oy	2761	CAGACAGTAAATATCACTGCAAGGCTTTCCTGTGGTGGTCAAGAAAGATPAAAGCCAGTTGAT	2820
Dp	2880	CAGACAGTAAATATCACTGCAAGGCTTTCCTGTGGTGGTCAAGAAAGATPAAAGCCAGTTGAT	2939
Oy	2821	AATGCCAAATGTAGTATCAAAAGAGAGGCTATAGTGTGTTGTCTATCATCTCAGTTCAAGAGC	2880
Dp	2940	AATGCCAAATGTAGTATCAAAAGAGAGGCTATAGTGTGTTGTCTATCATCTCAGTTCAAGAGC	2999
Oy	2881	AACGAAACCTGGACATTTACTCCAAATPAAACATGGACTTTTACAAAACCATATCGTATA	2940
Dp	3000	AACGAAACCTGGACATTTACTCCAAATPAAACATGGACTTTTACAAAACCATATCGTATA	3059
Oy	2941	CCACACCTTTTCCATCAAGTCACTTGGTTAAACTAAATGTAGAAAAATCTGCTAGAG	3000
Dp	3060	CCACACCTTTTCCATCAAGTCACTTGGTTAAACTAAATGTAGAAAAATCTGCTAGAG	3119
Oy	3001	GAAGAACTTTGAGAAACATTCATCAATGTCACTGAAAGAAATGGGAATGGAACATTTCCA	3060
Dp	3120	GAAGAACTTTGAGAAACATTCATCAATGTCACTGAAAGAAATGGGAATGGAACATTTCCA	3179
Oy	3061	AGTACAGTGAGCACAAATTAAGCCGTAATPAACTTATAGAGAAATGTTTTTAAAGAACCCAGC	3120
Dp	3180	AGTACAGTGAGCACAAATTAAGCCGTAATPAACTTATAGAGAAATGTTTTTAAAGAACCCAGC	3239
Oy	3121	TCAAGCAATATTAATGAAGTAAAGTTCCAGTATCTATGAAGTGGGCTCCAGTATTTATGA	3180
Dp	3240	TCAAGCAATATTAATGAAGTAAAGTTCCAGTATCTATGAAGTGGGCTCCAGTATTTATGA	3299

QY	3181	ATAGGCTTCCAGTGAATGAAGAAACATTTCAACACAGACTAGGTAAACACAGAGGCCAAATTTG	3240
Db	3300	ATAGGCTTCCAGTGAATGAAGAAACATTTCAACACAGACTAGGTAAACACAGAGGCCAAATTTG	3359
QY	3241	AATGCTATGCTTAGATTTAGGGGTTTTGCACCTGAGGCTATTAACAAGTCTCCGGA	3300
Db	3360	AATGCTATGCTTAGATTTAGGGGTTTTGCACCTGAGGCTATTAACAAGTCTCCGGA	3419
QY	3301	ACTAATTGTAGCATCTCGAANTAAAAAGCAAGATATGAGAAGTAGTTAGACTGTT	3360
Db	3420	AGTAATTTTAAGCATCTCGAANTAAAAAGCAAGATATGAGAAGTAGTTAGACTGTT	3479
QY	3361	AATACAGATTTCTCCATATCTGATTTCAATTAAGCTATGACAGGCTATGGGAAGT	3420
Db	3480	AATACAGATTTCTCCATATCTGATTTCAATTAAGCTATGACAGGCTATGGGAAGT	3539
QY	3421	CATGCATCTCAGGTTTTGTTCTGAGACACCTGATACCTGTTAGATGATGGTGAATAAG	3480
Db	3540	CATGCATCTCAGGTTTTGTTCTGAGACACCTGATACCTGTTAGATGATGGTGAATAAG	3599
QY	3481	GAAAGTACTAGTTTTGCTGGAATGACATTAGGAAGTTCTGCTGTTTTAGCAAAAGC	3540
Db	3600	GAAAGTACTAGTTTTGCTGGAATGACATTAGGAAGTTCTGCTGTTTTAGCAAAAGC	3659
QY	3541	GTCCAGAAAGGAGACTTACGAGAGTCTAGGCCCTTACACCCATACACTTTGGGCTCAG	3600
Db	3660	GTCCAGAAAGGAGACTTACGAGAGTCTAGGCCCTTACACCCATACACTTTGGGCTCAG	3719
QY	3601	GATTACCGAAGAGGGGCCAAGAATTAGAGTCTCAGAGAGACTTATCTAGTAGAGAT	3660
Db	3720	GATTACCGAAGAGGGGCCAAGAATTAGAGTCTCAGAGAGACTTATCTAGTAGAGAT	3779
QY	3661	GAAAGACTCTCCCTGCTTCAACACTGTTATTGTTAAAGTAACAATATACCTCTCAG	3720
Db	3780	GAAAGACTCTCCCTGCTTCAACACTGTTATTGTTAAAGTAACAATATACCTCTCAG	3839
QY	3721	TCTACTAGGCATAGCACCGTTGCTACGAGTGTCTGTAAAGAACACAGAGAAATTTA	3780
Db	3840	TCTACTAGGCATAGCACCGTTGCTACGAGTGTCTGTAAAGAACACAGAGAAATTTA	3899
QY	3781	TTATCATTTGAAGAATAGCTTAAATGACGTGCAATACAGGAATATTTGGCAAGGCATCT	3840
Db	3900	TTATCATTTGAAGAATAGCTTAAATGACGTGCAATACAGGAATATTTGGCAAGGCATCT	3959
QY	3841	CAGGAACATCACTTATAGTGAAGAAACAAATGTTCTGCTAGTGTGTTTCTTACAGTGC	3900
Db	3960	CAGGAACATCACTTATAGTGAAGAAACAAATGTTCTGCTAGTGTGTTTCTTACAGTGC	4019
QY	3901	AGTGAATTTGGAAGACTTAGTACGCAATATCAAAACCCAGAGTCTTTCTTGATTGGTTCT	3960
Db	4020	AGTGAATTTGGAAGACTTAGTACGCAATATCAAAACCCAGAGTCTTTCTTGATTGGTTCT	4079
QY	3961	TCCAACAACAAATGAGGCATCACTCTGAAAGCCAGGAGTGGTCTGAGTGAACGAATTTG	4020
Db	4080	TCCAACAACAAATGAGGCATCACTCTGAAAGCCAGGAGTGGTCTGAGTGAACGAATTTG	4139
QY	4021	GTTTCAGATGATGAAGAAAGAGAGCGGCTTGGAGAAGAAATATCAAGAAGCAAAAGC	4080
Db	4140	GTTTCAGATGATGAAGAAAGAGAGCGGCTTGGAGAAGAAATATCAAGAAGCAAAAGC	4199
QY	4081	ATGGAATTCAACTTAGGTGAAGCAGCACTGAGGAGTGAAGTGAATCAACACCGTCTGTGA	4140
Db	4200	ATGGAATTCAACTTAGGTGAAGCAGCACTGAGGAGTGAAGTGAATCAACACCGTCTGTGA	4259
QY	4141	GACTGCTCAGGGCTATCTCTCAGAGATGACATTTTAAACCTCAGCAGAGAGGATTAACATG	4200
Db	4260	GACTGCTCAGGGCTATCTCTCAGAGATGACATTTTAAACCTCAGCAGAGAGGATTAACATG	4319
QY	4201	CAACATTAAGCTGATTAAGCTTCACAGCAGAAATGGCTGAATAGAGCTGTGTTAGACAG	4260
Db	4320	CAACATTAAGCTGATTAAGCTTCACAGCAGAAATGGCTGAATAGAGCTGTGTTAGACAG	4379



OY	4261	ATATGGAGCCAGCCTTCTAACAGCTACCCCTCCATCTAAATGTCCTCTGGCCCTTGAG	4320
OY	4262	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	4321
Db	4380	CATGGAGGCCAGCCTTTTAAAGCTACCCCTTCCATCAATGTCCTTTCTGGCCCTTGAG	4439
OY	4321	GACCTGGAAATCCAGAAACCAAGCAGCATCGAAAAACAGTATTAACCTTCACAGAAAGT	4380
Db	4440	GACCTGGAAATCCAGAAACCAAGCAGCATCGAAAAACAGTATTAACCTTCACAGAAAGT	4499
OY	4381	AGTGAATACCTTAAAGCCAGAAATCCAGAAAGCCTTTCTGCTGCAAGTTTGAGGTGCT	4440
Db	4500	AGTGAATACCTTAAAGCCAGAAATCCAGAAAGCCTTTCTGCTGCAAGTTTGAGGTGCT	4559
OY	4441	GCAATATGTTCTACCAGTAAAAAATAAAGAACACAGAGTGGAAAGTATCCCTCTTAAA	4500
Db	4560	GCAATATGTTCTACCAGTAAAAAATAAAGAACACAGAGTGGAAAGTATCCCTCTTAAA	4619
OY	4501	TGCCATCATATGATGATGATGGTGTATCATATGCATGAGTTCCTGGAGCTCTCAATAAGA	4560
Db	4620	TGCCATCATATGATGATGATGGTGTATCATATGCATGAGTTCCTGGAGCTCTCAATAAGA	4679
OY	4561	AACATACCATCTCAAGAGAGAGCTCATTAAGGTTTGTATGTGAGAGCAACAGCTGGAA	4620
Db	4680	AACATACCATCTCAAGAGAGAGCTCATTAAGGTTTGTATGTGAGAGCAACAGCTGGAA	4739
OY	4621	GAGCTGTGGGCCACACGATTTGACGGAAAAACATCTTACTTGGCCAGGCAAGATCTAGAGGA	4680
Db	4740	GAGCTGTGGGCCACACGATTTGACGGAAAAACATCTTACTTGGCCAGGCAAGATCTAGAGGA	4799
OY	4681	ACCCCTTAACCTGSAATCTGSAATAGCCCTCTCTCATATGACCCCTGATCTGATCTCT	4740
Db	4800	ACCCCTTAACCTGSAATCTGSAATAGCCCTCTCTCATATGACCCCTGATCTGATCTCT	4859
OY	4741	GAAAGACAGAGCCCGAGAGTCAGCTGCTGTGGCAACATACCATCTTCAACCTCGCATTG	4800
Db	4860	GAAAGACAGAGCCCGAGAGTCAGCTGCTGTGGCAACATACCATCTTCAACCTCGCATTG	4919
OY	4801	AAATTTCCCCCAATGGAAGTTGCGAGATCTGCCCCAGAGTCCAGTGGCTGCTCAATACT	4860
Db	4920	AAATTTCCCCCAATGGAAGTTGCGAGATCTGCCCCAGAGTCCAGTGGCTGCTCAATACT	4979
OY	4861	GATACTGCTGGGTATTAATGCAATGGAAGAAAGTTGAGACAGGGAGAGCCAGATTTACA	4920
Db	4980	GATACTGCTGGGTATTAATGCAATGGAAGAAAGTTGAGACAGGGAGAGCCAGATTTACA	5039
OY	4921	GCTTCAACAGAAAGGTTCAACAAAAAGATGTCATGGTGTGTCGTGGCCTGACCCCGAAA	4980
Db	5040	GCTTCAACAGAAAGGTTCAACAAAAAGATGTCATGGTGTGTCGTGGCCTGACCCCGAAA	5099
OY	4981	GAATTTATGCTCGCTGTACAAAGTTTGGCCAGAAACACACATCACTTTAACTAATCTAAT	5040
Db	5100	GAATTTATGCTCGCTGTACAAAGTTTGGCCAGAAACACACATCACTTTAACTAATCTAAT	5159
OY	5041	ACTGAAGAGCATCTCATGTTGTTATGAAAAACAGATCTGTGTTGTGTGAAGGGACA	5100
Db	5160	ACTGAAGAGCATCTCATGTTGTTATGAAAAACAGATCTGTGTTGTGTGTGAAGGGACA	5219
OY	5101	CTGAATATATTTTCTAGGAATTTGCGGAGAGAAATGGGTAGTTAGCTATTTCTGGGTAC	5160
Db	5220	CTGAATATATTTTCTAGGAATTTGCGGAGAGAAATGGGTAGTTAGCTATTTCTGGGTAC	5279
OY	5161	CAGCTCTATTAAAGAAAAAAATGCTAATGAGCATATTTTGAGTGCAGAGAGATGTG	5220
Db	5280	CAGCTCTATTAAAGAAAAAAATGCTAATGAGCATATTTTGAGTGCAGAGAGATGTG	5339
OY	5221	GTCATATGGAAGAAACCCCAAGGTCCTAAAGGACAGAGATCCCGAGGACAGAAAGATC	5280
Db	5340	GTCATATGGAAGAAACCCCAAGGTCCTAAAGGACAGAGATCCCGAGGACAGAAAGATC	5399
OY	5281	TTTCAAGGGGGCTAGAAATCTGTGCTATGAGGCCCTTCCACCAATGCCACAGATCAACTG	5340
Db	5400	TTTCAAGGGGGCTAGAAATCTGTGCTATGAGGCCCTTCCACCAATGCCACAGATCAACTG	5459
OY	5341	GAATGATTCGTACACGCTGTGTGCTGCTTGTGTGTGAAGGACCTTTCATATTCACCTTT	5400

Dd		5460	GAAATGATAGTACAGCTGTGTGGTGCTTCCTGTGGTGAAGAAGCTTTCACTCATTACCCCTT	5519
Oy		5401	GGCAGAGGTGCCAACCATAATTTGGTTGTGTGCAGCCAGATGCGCTGGACAGACAAATGGC	5460
Dd		5520	GGCACAGGTGTCCAACCCAATTTGTGGTTGTGCAGCCAGATGCGCTGGACAGACAAATGGC	5579
Oy		5461	TTCCATCATTAATGGGGCAGATGTGTGAGGCACCTTGTGTGACCCGACAGTGGGTGTGGAC	5520
Dd		5580	TTCCATCATTAATGGGGCAGATGTGTGAGGCACCTTGTGTGACCCGACAGTGGGTGTGGAC	5639
Oy		5521	AGTGTAGACTCTTACCAAGTGGCCAGAGACTGGACACCCTACCTGATACCCCAGATCCCCCACC	5580
Dd		5640	AGTGTAGACTCTTACCAAGTGGCCAGAGACTGGACACCCTACCTGATACCCCAGATCCCCCACC	5699
Oy		5581	AGCCACTTAC 5589	
Dd		5700	AGCCACTTAC 5708	
	RESULT 10			
	LOCUS	AR008159	5914 bp DNA Linear PAT 04-DEC-1998	
	DEFINITION	Sequence 1 from patent US 5753441.		
	ACCESSION	AR008159		
	VERSION	AR008159.1 GI:3967268		
	KEYWORDS	.		
	SOURCE	Unknown.		
	ORGANISM	Unknown.		
	REFERENCE	Unclassified.		
	AUTHORS	1 (bases 1 to 5914)		
		Skolnick,M.H., Goldgar,D.E., Miki,Y., Swenson,J., Kamb,A.,		
		Hartshman,K.D., Shattuck-Eidens,D.M., Tavtiglian,S.V., Wiseman,R.W.		
		and Futreal,P.Andrew.		
	TITLE	170-linked breast and ovarian cancer susceptibility gene		
	JOURNAL	Patent: US 5753441-A 1 19-MAY-1998;		
	FEATURES	Location/Qualifiers		
	Source	1..5914		
		/organism="unknown"		
	BASE COUNT	2006 a 1156 c 1316 g 1436 t		
	ORIGIN			
	Query Match	99.9%; Score 5585.8; DB 6; length 5914;		
	Best Local Similarity	100.0%; Pred. No. 0;		
	Matches 5587; Conservative	0; Mismatches 2; Indels 0; Gaps 0;		
Oy		1 ATGATATTATCTGCTCTTGGCGTTGAGAAAGTACAAAATGTCATTATGCTATGACAAAA	60	
Dd		120 ATGGAATTAATCTGCTCTTGGCGTTGAGAAAGTACAAAATGTCATTATGCTATGACAAAA	179	
Oy		61 ATTCTTAGAGTGTCCCATCTGTCTGGAGTTGATCAGGAACCTGTCTCACAAAGTGTGAC	120	
Dd		180 ATCTTAGAGTGTCCCATCTGTCTGGAGTTGATCAGGAACCTGTCTCACAAAGTGTGAC	239	
Oy		121 CACATATTTTTGCCAATTTTGCATGCTGAACCTTCACAACAGAAAGAGGCTTCACAG	180	
Dd		240 CACATATTTTTGCCAATTTTGCATGCTGAACCTTCACAACAGAAAGAGGCTTCACAG	289	
Oy		181 TGTCCTTTAGTAGANTGATNTATACAAAAGAGCCTTACAAGAAATATAGACATTAGT	240	
Dd		300 TGTCTCTTTAGTAGAATGATNTATACAAAAGAGCCTTACAAGAAATATAGACATTAGT	359	
Oy		241 CAACCTTTTGAAGAGCATTTGAAANAATCATTTGTCTTTCAGCTTGACAGAGTTTGGAG	300	
Dd		360 CAACCTTTTGAAGAGCATTTGAAANAATCATTTGTCTTTCAGCTTGACAGAGTTTGGAG	419	
Oy		301 TAGCAACAACAGCTATTAATTTTGCAAAAAAGAAANAATACCTCTGTAACATCTAANAAGAT	360	
Dd		420 TAGCAACAACAGCTATTAATTTTGCAAAAAAGAAANAATACCTCTGTAACATCTAANAAGAT	479	
Oy		361 GAATTTTCTATCATCCCAAATGATGGGCTACAGAAACCGTGCACAAAAGACTTCTACAGAGT	420	
Dd		480 GAATTTTCTATCATCCCAAATGATGGGCTACAGAAACCGTGCACAAAAGACTTCTACAGAGT	539	



421 GAACCCGAAATCTCTTCTTGAGAAACAGTCTCAGTCTCCACTCTCTAACCTTGA 480  
|||||  
540 GAACCCGAAATCTCTTCTTGAGAAACAGTCTCAGTCTCCACTCTCTAACCTTGA 539  
481 ACTGTGAGACCTCTGAGGACAAAGCAGCGGATACAACTCTAAAGAGCTCTCTACAT 540  
600 ACTGTGAGACCTCTGAGGACAAAGCAGCGGATACAACTCTAAAGAGCTCTCTACAT 659  
541 GAATGGGATCTGATCTCTGAGATACCGTTAATAGGCACTAATTCAGCTGAGGGA 600  
660 GAATGGGATCTGATCTCTGAGATACCGTTAATAGGCACTAATTCAGCTGAGGGA 719  
601 GATCAAGATTTGTACAAATCAACCCCTCAAGGACAGGAGATGAATCAGTTGATTC 650  
720 GATCAAGATTTGTACAAATCAACCCCTCAAGGACAGGAGATGAATCAGTTGATTC 779  
661 GCAAAAAAGCGCTGCTGTGTAATTTCTGAGACGATGTAACAATCTGACATCA 720  
780 GCAAAAAAGCGCTGCTGTGTAATTTCTGAGACGATGTAACAATCTGACATCA 839  
721 CCCAGTAATGATTTGAACACCACTGAGAGCGTGAAGGCTGAGAGGCTCCAGAAAG 780  
840 CCCAGTAATGATTTGAACACCACTGAGAGCGTGAAGGCTGAGAGGCTCCAGAAAG 839  
781 TATCAGGGTATGTTCTGTTTCAAACTTGATGAGCCATGTGGCAAAATCTCATGCC 840  
900 TATCAGGGTATGTTCTGTTTCAAACTTGATGAGCCATGTGGCAAAATCTCATGCC 959  
841 AGCTATTACAGATGAGACAGCAGCTTTATTACTACTAAGACAGATGAATGTGA 900  
960 AGCTATTACAGATGAGACAGCAGCTTTATTACTACTAAGACAGATGAATGTGA 1019  
901 AAGGCTGAATTTGTATATAAAGCAACAGCGCTGTTAGCAAGGAGCCAACTAACA 960  
1020 AAGGCTGAATTTGTATATAAAGCAACAGCGCTGTTAGCAAGGAGCCAACTAACA 1079  
961 TGGGCTGGAAGTAAAGAAACATGTAATGATAGGCGGACTCCAGCAGAAAAAAGTA 1020  
1080 TGGGCTGGAAGTAAAGAAACATGTAATGATAGGCGGACTCCAGCAGAAAAAAGTA 1139  
1021 GATCTGATGCTGATCCCTGTGTGAGAAAAAAGATGCAATTAAGCAGAACTCCATGC 1080  
1140 GATCTGATGCTGATCCCTGTGTGAGAAAAAAGATGCAATTAAGCAGAACTCCATGC 1199  
1081 TCAGAGAACTCTAGAGATCTAGAGATGTTCTTGATTAACACTAATAGCAGATTGAG 1140  
1200 TCAGAGAACTCTAGAGATCTAGAGATGTTCTTGATTAACACTAATAGCAGATTGAG 1259  
1141 AAAGTTAATGAGTGTCTTCCAGAGTGATGAACCTGTAGGTTCTGATGACTCATGAT 1200  
1260 AAAGTTAATGAGTGTCTTCCAGAGTGATGAACCTGTAGGTTCTGATGACTCATGAT 1319  
1201 GGGGAGTCTGATCAAAATGCCAAAGTGTATTTGAGCGTTCTTAATAGTATGAT 1260  
1330 GGGGAGTCTGATCAAAATGCCAAAGTGTATTTGAGCGTTCTTAATAGTATGAT 1379  
1261 GAATATTCCTGTTCTTCCAGAAAAATAGACTACTGGCCAGTATCCTATGAGGCTTGA 1320  
1380 GAATATTCCTGTTCTTCCAGAAAAATAGACTACTGGCCAGTATCCTATGAGGCTTGA 1439  
1321 ATATGTAAAGTGAAGAAGTCTCAATCAATGAGAGATGAATTTAGACAAATAA 1380  
1440 ATATGTAAAGTGAAGAAGTCTCAATCAATGAGAGATGAATTTAGACAAATAA 1499  
1381 TTTGGGAAAACCTATCGGAGAGAGCAGCCCTCCCACTTAAGCCATGTAATCTGAAAT 1440  
1500 TTTGGGAAAACCTATCGGAGAGAGCAGCCCTCCCACTTAAGCCATGTAATCTGAAAT 1559  
1441 CTAATTTATAGGAGCTTTGTACTGAGCAGAGATATACAGAGGCTCCCTCACAAT 1500  
1560 CTAATTTATAGGAGCTTTGTACTGAGCAGAGATATACAGAGGCTCCCTCACAAT 1619

1501 AATTTAAGCGTAAAGAGAGACCTACATCAGGCTTCATCTGAGGATTTTATCAGAAA 1560  
|||||  
1620 AATTTAAGCGTAAAGAGAGACCTACATCAGGCTTCATCTGAGGATTTTATCAGAAA 1679  
1561 GCAGATTTGGCAGTTTAAAGAGCTCCTGAATGATTAATCAGGGAACCTAACAGGAG 1620  
1680 GCAGATTTGGCAGTTTAAAGAGCTCCTGAATGATTAATCAGGGAACCTAACAGGAG 1739  
1621 CAGATGTCAGTGAATGATTAATTAATAGGTCATGAGGATTAATTAACAAAGGTAT 1680  
1740 CAGATGTCAGTGAATGATTAATTAATAGGTCATGAGGATTAATTAACAAAGGTAT 1799  
1681 TCTATTGAGATGAGAAAAATCTTAACCAATAGATCACTGAAAAAGATCTCTTC 1740  
1800 TCTATTGAGATGAGAAAAATCTTAACCAATAGATCACTGAAAAAGATCTCTTC 1859  
1741 AAACGAAAGCTGACCTATTAAGCAGCAGTATTAAGCAATATGGAATCGAATTAATATC 1800  
1860 AAACGAAAGCTGACCTATTAAGCAGCAGTATTAAGCAATATGGAATCGAATTAATATC 1919  
1801 CACATTTCAAAAGCAGCTTAAAGAAAGATAGGCTGAGAGGAGTCTTCAACGAGCATAT 1860  
1920 CACATTTCAAAAGCAGCTTAAAGAAAGATAGGCTGAGAGGAGATCTTCAACGAGCATAT 1979  
1861 CATGCGCTTGAACCTAGTACGTAGTAATCTAAGCCCACTAATTTGATCTGAATTCGA 1920  
1980 CATGCGCTTGAACCTAGTACGTAGTAATCTAAGCCCACTAATTTGATCTGAATTCGA 2039  
1921 ATTGATAGTGTGTTCTGACAGTGAAGATTAAGAAAAAAGTATACACCAGTATG 1980  
2040 ATTGATAGTGTGTTCTGACAGTGAAGATTAAGAAAAAAGTATACACCAGTATG 2099  
1981 AGGACAGAGAAACCTACACATCATGAGGTTAAAGAACTGCAATGAGAGCCAAAG 2040  
2100 AGGACAGAGAAACCTACACATCATGAGGTTAAAGAACTGCAATGAGAGCCAAAG 2159  
2041 AGTAACAAAGCCAAATGACAGACAAATGAAGACATGACAGTATCTTCCAGAGCTG 2100  
2160 AGTAACAAAGCCAAATGACAGACAAATGAAGACATGACAGTATCTTCCAGAGCTG 2219  
2101 AAGTTAACAATGACACCTGTTCTTTACTAAGTGTCAATTAACGAGTAATTAAGAA 2160  
2220 AAGTTAACAATGACACCTGTTCTTTACTAAGTGTCAATTAACGAGTAATTAAGAA 2279  
2161 TTTGCAATCTAGCCCTTCCAGAGAGAAAAAGAAAGTGAAGTGAAGTGAAGTGAAG 2220  
2280 TTTGCAATCTAGCCCTTCCAGAGAGAAAAAGAAAGTGAAGTGAAGTGAAGTGAAG 2339  
2221 TCTAATATGCTGAAGACCCCAAGATCATGTTAAGTGAAGAAAGGTTTGCACACT 2280  
2340 TCTAATATGCTGAAGACCCCAAGATCATGTTAAGTGAAGAAAGGTTTGCACACT 2399  
2281 GAAAGATCTGTAGAGATGACGATATTTGATGATGATGATGATGATGATGATGATG 2340  
2400 GAAAGATCTGTAGAGATGACGATATTTGATGATGATGATGATGATGATGATGATG 2459  
2341 GAAAGATCTGTAGAGATGACGATATTTGATGATGATGATGATGATGATGATGATG 2400  
2460 GAAAGATCTGTAGAGATGACGATATTTGATGATGATGATGATGATGATGATGATG 2519  
2401 TGTGTAGTCAAGTGTGACAGATTTGAAAAACCCCAAGGAGTAAATTCATGTTTCCAAA 2460  
2520 TGTGTAGTCAAGTGTGACAGATTTGAAAAACCCCAAGGAGTAAATTCATGTTTCCAAA 2579  
2461 GATATATGAATGACAGAGAGGCTTTAAGTATCATTTGGGACATGAAGTTAACACAGT 2520  
2580 GATATATGAATGACAGAGAGGCTTTAAGTATCATTTGGGACATGAAGTTAACACAGT 2639  
2521 CGGGAACAGCATGGAATGAGAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2580  
2640 CGGGAACAGCATGGAATGAGAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2639  
2581 TTCAAGCTTTCAAGGCGCAGTCAATTTGCTGTTTCAAAATCCAGAAATGACAGAAAG 2640

|||||  
Db 2700 TTCAAGTTTCAAGGCCAGCATTTGCTCCGTTTCAATCAGAGAAATGAGAG 2759  
QY 2641 GAATGCGCAACATTTCTGCCCACCTCGGGCTCTTAAAGAAACAAAGTCACT 2700  
Db 2760 GAATGCGCAACATTTCTGCCCACCTCGGGCTCTTAAAGAAACAAAGTCACT 2819  
QY 2701 TTGAATGTAGCAAAAGAAAGAAATCAAGAAAGAAATGATGATTAATCAAGCCCTGTA 2760  
Db 2820 TTGTAATGTAGCAAAAGAAAGAAATCAAGAAAGAAATGATGATTAATCAAGCCCTGTA 2879  
QY 2761 CAGACATTAATATCACTGAGGCTTCTGTTGGTGTGAGAAAGTAAAGCCAGTTGAT 2820  
Db 2880 CAGACATTAATATCACTGAGGCTTCTGTTGGTGTGAGAAAGTAAAGCCAGTTGAT 2939  
QY 2821 AATGCCAAATGTAGTATCAAGAGAGGCTAGGTTTGTCTATCATCTGCTCAGAGC 2880  
Db 2940 AATGCCAAATGTAGTATCAAGAGAGGCTAGGTTTGTCTATCATCTGCTCAGAGC 2999  
QY 2881 AAGGAACTGAGCTATTAATCAATTAATGAGACTTTTACAAACCCATATCTGATA 2940  
Db 3000 AAGGAACTGAGCTATTAATCAATTAATGAGACTTTTACAAACCCATATCTGATA 3059  
QY 2941 CCACCACTTTTCCCATCAAGTCATTTGTTAAATGTAAGAAATCTGCTAGAG 3000  
Db 3060 CCACCACTTTTCCCATCAAGTCATTTGTTAAATGTAAGAAATCTGCTAGAG 3119  
QY 3001 GAAACCTTTAGAGACATTAATGATGACCTGGAAGAGAAATGGGAATGGAACATTTCA 3060  
Db 3120 GAAACCTTTAGAGACATTAATGATGACCTGGAAGAGAAATGGGAATGGAACATTTCA 3179  
QY 3061 AGTACAGTGAAGCAATTAATGATGACCTGGAAGAGAAATGGGAATGGAACATTTCA 3120  
Db 3180 AGTACAGTGAAGCAATTAATGATGACCTGGAAGAGAAATGGGAATGGAACATTTCA 3239  
QY 3121 TCAGCAATATTAATGATGAGTGTCCAGTACTAATGAGTGGGCTCCAGTATTAATGAA 3180  
Db 3240 TCAGCAATATTAATGATGAGTGTCCAGTACTAATGAGTGGGCTCCAGTATTAATGAA 3299  
QY 3181 ATAGGTTCCAGTATGAAATCAATTCAGAGAACTAGTGAACAGAGGCCCAAAATG 3240  
Db 3300 ATAGGTTCCAGTATGAAATCAATTCAGAGAACTAGTGAACAGAGGCCCAAAATG 3359  
QY 3241 AATGCTATGCTTATGATGAGGTTTGGCAACTGAGGCTATTAACAAATCTTCTGGA 3300  
Db 3360 AATGCTATGCTTATGATGAGGTTTGGCAACTGAGGCTATTAACAAATCTTCTGGA 3419  
QY 3301 AGTAAATGTAGCAATCTGGAATTAAGAAAGCAAGATATGAGAGAGTTCAGACTGT 3360  
Db 3420 AGTAAATGTAGCAATCTGGAATTAAGAAAGCAAGATATGAGAGAGTTCAGACTGT 3479  
QY 3361 AATACGATTTCTCTCATATCTGATTTTCAGATTAATTAAGAGCTATGGAAGTACT 3420  
Db 3480 AATACGATTTCTCTCATATCTGATTTTCAGATTAATTAAGAGCTATGGAAGTACT 3539  
QY 3421 CATGCAATCTGAGTTTGTCTGAGACACCTGATGACTGTTAGATGATGTTGGAATTAAG 3480  
Db 3540 CATGCAATCTGAGTTTGTCTGAGACACCTGATGACTGTTAGATGATGTTGGAATTAAG 3599  
QY 3481 GAAGATATAGTATTTGCTGGAATTAAGCAATTAAGCAATTTCTGTTTTCAGAAAGC 3540  
Db 3600 GAAGATATAGTATTTGCTGGAATTAAGCAATTAAGCAATTTCTGTTTTCAGAAAGC 3659  
QY 3541 GTCCAGAAAGAGAGCTTAGCAGAGGCTTACGCCCTTTTCACCAATACACTTTGGCTCAG 3600  
Db 3660 GTCCAGAAAGAGAGCTTAGCAGAGGCTTACGCCCTTTTCACCAATACACTTTGGCTCAG 3719  
QY 3601 GGTACCGAAGAGGCGCAAGAAATTAAGATCTCCAGAAAGAACTATATAGAGAGAT 3660  
Db 3720 GGTACCGAAGAGGCGCAAGAAATTAAGATCTCCAGAAAGAACTATATAGAGAGAT 3779  
QY 3661 GAAGAGCTTCCGCTTCCAACTGTTATTTGTAAGTAAACATATACCTTCTCAG 3720  
|||||

Db 3780 GAAGAGCTTCCGCTTCCAACTGTTATTTGTAAGTAAACAAATATACCTTCTCAG 3839  
QY 3721 TCTACTAGCATATGACAGGCTTGTACAGAGTGTCTTAAGAACAGAGAGAAATTTA 3780  
Db 3840 TCTACTAGCATATGACAGGCTTGTACAGAGTGTCTTAAGAACAGAGAGAAATTTA 3899  
QY 3781 TTATCATTTGAAGATATGATTAATGATGAGTAACTGAGTAAATATGAGCAAGATCT 3840  
Db 3900 TTATCATTTGAAGATATGATTAATGATGAGTAACTGAGTAAATATGAGCAAGATCT 3959  
QY 3841 CAGAACATCACTTATGAGAGAAACAAATGTTGCTGATGCTGTTTCTTACAGTGC 3900  
Db 3960 CAGAACATCACTTATGAGAGAAACAAATGTTGCTGATGCTGTTTCTTACAGTGC 4019  
QY 3901 AGTGAATTTGGAAGACTGACTGCAATATCAACACCAGATCTCTTCTGATTTGTTCT 3960  
Db 4020 AGTGAATTTGGAAGACTGACTGCAATATCAACACCAGATCTCTTCTGATTTGTTCT 4079  
QY 3961 TCCAAACAAATGAGGCTATGATGAGAGGAGGAGTGTGCTGAGTGAAGAGATTC 4020  
Db 4080 TCCAAACAAATGAGGCTATGATGAGAGGAGGAGTGTGCTGAGTGAAGAGATTC 4139  
QY 4021 GTTTCAGATGATGAGAGAGAGAGAGGCTTGGAGAAATATCAAGAGAGCAAGC 4080  
Db 4140 GTTTCAGATGATGAGAGAGAGAGAGGCTTGGAGAAATATCAAGAGAGCAAGC 4199  
QY 4081 ATGATTTCAACTTATGATGAGAGAGAGAGGCTTGGAGAAATATCAAGAGAGCAAGC 4140  
Db 4200 ATGATTTCAACTTATGATGAGAGAGAGAGGCTTGGAGAAATATCAAGAGAGCAAGC 4259  
QY 4141 GACTGCTCAGGCTATGCTCTCAGAGTATCTTAACTGAGTGAAGAGAGATACCATG 4200  
Db 4260 GACTGCTCAGGCTATGCTCTCAGAGTATCTTAACTGAGTGAAGAGAGATACCATG 4319  
QY 4201 CAACATTAATCTGATTAAGCTCCAGACAGAAATGGCTGAGTGAAGCTGTGTTAGACAG 4260  
Db 4320 CAACATTAATCTGATTAAGCTCCAGACAGAAATGGCTGAGTGAAGCTGTGTTAGACAG 4379  
QY 4261 CATGGAGGAGGCTTCTAAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAG 4320  
Db 4380 CATGGAGGAGGCTTCTAAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAG 4439  
QY 4321 GACCTGCGAATCCAGAACACACATVCAAGAAACAGTATTAATTCACAGAAAGT 4380  
Db 4440 GACCTGCGAATCCAGAACACACATVCAAGAAACAGTATTAATTCACAGAAAGT 4499  
QY 4381 AGTGAATATCCCTATTAAGCCAGAAATCCAGAGGCTTCTGTCAGAGTGTGCT 4440  
Db 4500 AGTGAATATCCCTATTAAGCCAGAAATCCAGAGGCTTCTGTCAGAGTGTGCT 4559  
QY 4441 GCAGATATGTTCTACAGTAAATTAAGAACACAGAGTGAAGAGTATCCCTTCTAAA 4500  
Db 4560 GCAGATATGTTCTACAGTAAATTAAGAACACAGAGTGAAGAGTATCCCTTCTAAA 4619  
QY 4501 TGCCCATCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4560  
Db 4620 TGCCCATCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4679  
QY 4561 AACTACCATCTCAAGAGAGCTATTAAGGTTGTTGATGTGAGAGAGCAACAGCTGAA 4620  
Db 4680 AACTACCATCTCAAGAGAGCTATTAAGGTTGTTGATGTGAGAGAGCAACAGCTGAA 4739  
QY 4621 GAGTCTGGGCGCAACGATTTGAGGAAACATCTTACTTGGCAAGGCAAGATCTAAGAGGA 4680  
Db 4740 GAGTCTGGGCGCAACGATTTGAGGAAACATCTTACTTGGCAAGGCAAGATCTAAGAGGA 4799  
QY 4681 ACCCTTACCTGGAATCTGGAATCAGGCTCTCTCTGATATACCTGATCTGATCTCTCT 4740  
Db 4800 ACCCTTACCTGGAATCTGGAATCAGGCTCTCTCTGATATACCTGATCTGATCTCTCTCT 4859  
QY 4741 GAAGACAGAGCCCGAGAGTCTGTTGGCAACATATCACTTCAACCTCTGCAATG 4800  
Db 4860 GAAGACAGAGCCCGAGAGTCTGTTGGCAACATATCACTTCAACCTCTGCAATG 4919  
|||||

OY	4801	AAAGTTCCTCCCAATGGAAGTGTGCAAGAACTCGCCAGAGATCCAGCTGCCTGCATACACT	4860
Db	4920	AAAGTTCCTCCCAATGGAAGTGTGCAAGAACTCGCCAGAGATCCAGCTGCCTGCATACACT	4979
OY	4861	GATACCTCGTGGGTATATATGCAATGTGAAGAGTGTGAGCAGAGGAGAAAGCCAGATTGACA	4920
Db	4980	GATACCTCGTGGGTATATATGCAATGTGAAGAGTGTGAGCAGAGGAGAAAGCCAGATTGACA	5039
OY	4921	GCTTCACAGAAAGGGTCAACAAAAAGATGTCCATGGTGGTGTGCTGGCTTGACCCAGAA	4980
Db	5040	GCTTCACAGAAAGGGTCAACAAAAAGATGTCCATGGTGGTGTGCTGGCTTGACCCAGAA	5099
OY	4981	GAATTATATGTCGCTGTACAGTTGGCCAGAAACCCGATCAGTTAACTAACTAATT	5040
Db	5100	GAATTATATGTCGCTGTGTACAGTTGGCCAGAAACCCGATCAGTTAACTAACTAATT	5159
OY	5041	ACTGAAGAGACTACTCATGTGTATATGAAAAACAGATGCTGAGTTGTGTGTGAACGGACA	5100
Db	5160	ACTGAAGAGACTACTCATGTGTATATGAAAAACAGATGCTGAGTTGTGTGTGAACGGACA	5219
OY	5101	CTGAATAATTTTCTTAGCAATTGCGGGAGGAAAAATGGTATGTTACCTATTTCTGGGTACC	5160
Db	5220	CTGAATAATTTTCTTAGCAATAATGCGGGAGGAAAAATGGTATGTTACCTATTTCTGGGTACC	5279
OY	5161	CAGCTATATTAAGAAAGAAAAATGCTGAATGAGCATGTTTGAAGTCAGAGGAGATGTG	5220
Db	5280	CAGCTATATTAAGAAAGAAAAATGCTGAATGAGCATGTTTGAAGTCAGAGGAGATGTG	5339
OY	5221	GTCAATGGAAGAAACACCAAGGTCCTCAAGGAGCAGCAAGAAATCCACAGACAGAAAGATC	5280
Db	5340	GTCAATGGAAGAAACACCAAGGTCCTCAAGGAGCAGCAAGAAATCCACAGACAGAAAGATC	5399
OY	5281	TTTCAGGGGGCTAGAAATCTGTTGCTATGAGGGCCCTTCCAAATGAGCCACAGATCACTG	5340
Db	5400	TTTCAGGGGGCTAGAAATCTGTTGCTATGAGGGCCCTTCCAAATGAGCCACAGATCACTG	5459
OY	5341	GAATGGAATGTACAGCTGTGTGTGCTTCTGTGTGAAGAGCACTTTCATCTTACCCCTT	5400
Db	5460	GAATGGAATGTACAGCTGTGTGTGCTTCTGTGTGAAGAGCACTTTCATCTTACCCCTT	5519
OY	5401	GGCACAGGTGTCCACCCCAATTGTGTGTGTGACCAAGATGCTGTGACAGAGACAAATGAC	5460
Db	5520	GGCACAGGTGTCCACCCCAATTGTGTGTGTGACCAAGATGCTGTGACAGAGAGACAAATGAC	5579
OY	5461	TTTCATGCAATTGGGGCAGATGTGTGAGGCACCTGTGGTGGACCCGAGAGTGGGTGTGAGC	5520
Db	5580	TTTCATGCAATTGGGGCAGATGTGTGAGGCACCTGTGGTGGACCCGAGAGTGGGTGTGAGC	5639
OY	5521	AGTGTAGCACTCTACAGTGTGACAGAGCTGAGACCTACCTGATATACCCAGATCCCCAC	5580
Db	5640	AGTGTAGCACTCTACAGTGTGACAGAGCTGAGACCTACCTGATATACCCAGATCCCCAC	5699
OY	5581	AGGCACCTAC 5589	
Db	5700	AGGCACCTAC 5708	

RESULT 11				
LOCUS	AR136942	5914 bp	DNA	linear
DEFINITION	Sequence 1 from patent US 6162897.			
ACCESSION	AR136942			
VERSION	AR136942.1	GI:14478192		
KEYWORDS				
SOURCE	Unknown.			
ORGANISM	Unknown.			
REFERENCE	Unclassified.			
AUTHORS	1 (bases 1 to 5914)			
	Skolnick,M.H., Goldgar,D.E., Miki,Y., Swenson,J., Kamb,A.,			
	Hershman,K.D., Shattuck-Eidens,D.M., Tavtigian,S.V., Wiseman,R.W.			
	and Fureted,P.Andrew.			
TITLE	17q-linked breast and ovarian cancer susceptibility gene			

JOURNAL	Patent: US 6162897-A 1 19-DEC-2000
FEATURES	Location/Qualifiers
source	1. .5914
	/Organization="unknown"
BASE COUNT	2006 a 1156 c 1316 g 1436
ORIGIN	

Query Match	99.9%	Score 5585.8	DB 6	Length 5914
Best Local Similarity	100.0%	Pred. No. 0		
Matches 5587: Conservative	0	Mismatches 2	Indels 0	Gaps 0

Oy	ATGATATTATCTCTCTTCGCGTTGAGAAAGTACAAATATGCATTAATGCTATGCGAGAAA	60
Db	120 ATGATATTATCTCTCTTCGCGTTGAGAAAGTACAAATATGCATTAATGCTATGCGAGAAA	179
Oy	61 ATCTTAGAGTGTCCCATCTCTCGAGATTGATCAAGAACTGTCTCCACAAAGTGTAC	120
Db	180 ATCTTAGAGTGTCCCATCTCTCGAGATTGATCAAGAACTGTCTCCACAAAGTGTAC	239
Oy	121 CACATATTTTGGCAATTTTGCATGCTGGAACCTTCTCAACCAGAAGAAAGGCGCTTCACAG	180
Db	240 CACATATTTTGGCAATTTTGCATGCTGGAACCTTCTCAACCAGAAGAAAGGCGCTTCACAG	299
Oy	181 TGTCTTATGTAGAATATATTAACCCAAAAGAGCCATCAGAAAGATATGAGATTAGT	240
Db	300 TGTCTTATGTAGAATATATTAACCCAAAAGAGCCATCAGAAAGATATGAGATTAGT	359
Oy	241 CAACCTGTTCAGAGCTATTTGAAATCATTTCTGCTTTACGTTGACACAGGTTTGAG	300
Db	360 CAACCTGTTCAGAGCTATTTGAAATCATTTCTGCTTTACGTTGACACAGGTTTGAG	419
Oy	301 TATGCAAAACGCTATATTTTGCACAAAAGGAAATTAATCTCTGTAACCTCTAAAGAT	360
Db	420 TATGCAAAACGCTATATTTTGCACAAAAGGAAATTAATCTCTGTAACCTCTAAAGAT	479
Oy	361 GAAGTTCTATCATCCCAAGGTATGGGCTACAGAAACCGTCGCCAAAAGACTTTCACAGAGT	420
Db	480 GAAGTTCTATCATCCCAAGGTATGGGCTACAGAAACCGTCGCCAAAAGACTTTCACAGAGT	539
Oy	421 GAACCCGAAATATCTCTTCTTGGCAGGAACACAGTCTCAGTGTCCAACTCTTAACCTTTGGA	480
Db	540 GAACCCGAAATATCTCTTCTTGGCAGGAACACAGTCTCAGTGTCCAACTCTTAACCTTTGGA	599
Oy	481 ACAGTGAACCTGTGAGCAAGGACAGGGATACACCTCCAAAAGAGCTGTGCTCATTT	540
Db	600 ACAGTGAACCTGTGAGCAAGGACAGGGATACACCTCCAAAAGAGCTGTGCTCATTT	659
Oy	541 GAAATGGGATCTGATTTCTTCTGAGATACCGTTAATAAGGCACTTTATTCAGTGTGGGA	600
Db	660 GAAATGGGATCTGATTTCTTCTGAGATACCGTTAATAAGGCACTTTATTCAGTGTGGGA	719
Oy	601 GATCAGAGATTTGTTACAAATTCACCCCTCAAGGAACAGAGGATGAAATCATGTTGGATTCT	660
Db	720 GATCAGAGATTTGTTACAAATTCACCCCTCAAGGAACAGAGGATGAAATCATGTTGGATTCT	779
Oy	661 GCAAAAAAGGCTCTTCTGAAATTTTCTGAGACGATGTAAACAATFACTGAAATCATCAA	720
Db	780 GCAAAAAAGGCTCTTCTGAAATTTTCTGAGACGATGTAAACAATFACTGAAATCATCAA	839
Oy	721 CCCAGTAATATGATTTGAACACCACTGAGAGAGCGTGCAGTGAAGGATCCAGAAAAG	780
Db	840 CCCAGTAATATGATTTGAACACCACTGAGAGAGCGTGCAGTGAAGGATCCAGAAAAG	899
Oy	781 TATCAGAGGATGCTGTTTCAAACTTGTGATGAGACATGTGGCCAAATATCTCATGGCC	840
Db	900 TATCAGAGGATGCTGTTTCAAACTTGTGATGAGACATGTGGCCAAATATCTCATGGCC	959
Oy	841 AGCTCATTTACAGCAAGACAGCAGTTATTACTCTAAAGACAGAAATGATGTAGAA	900
Db	960 AGCTCATTTACAGCAAGACAGCAGTTATTACTCTAAAGACAGAAATGATGTAGAA	1011
Oy	901 AAGGCTGAATCTGTATATAAAGCAACAGCCTGCTTAAAGAGACCAATATAACAGA	960

Db 1020 AAGGCTGAATTCGTGAATTAAGCAACAGCTGGCTTAGCCAGGACCAACATTAACAGA 1079  
Qy TGGGCTGGAAGTAAGAAACATGTAATGATAGGGGAGCTCCAGCACAGAAAAAAGSTA 1020  
Db TGGGCTGGAAGTAAGAAACATGTAATGATAGGGGAGCTCCAGCACAGAAAAAAGSTA 1139  
Qy 1021 GATCTGAATGCTGATCCCTGTGTGTGAGAGAAAAAGATGAATAGCAGAAACCTGCCATGC 1080  
Db 1140 GATCTGAATGCTGATCCCTGTGTGTGAGAGAAAAAGATGAATAGCAGAAACCTGCCATGC 1139  
Qy 1081 TCAGAGAACTCTAGAGATGAGAGATGTTCTTGATTAACACTAAATAGCAGCATTCAG 1140  
Db 1200 TCAGAGAACTCTAGAGATGAGAGATGTTCTTGATTAACACTAAATAGCAGCATTCAG 1259  
Qy 1141 AAGATTAATGAGTGGTTTCCAGAAAGTGAATGATGATGTTAGTGTGATGACTGCATGAT 1280  
Db 1260 AAGATTAATGAGTGGTTTCCAGAAAGTGAATGATGATGTTAGTGTGATGACTGCATGAT 1319  
Qy 1201 GGGGAGTCTGAATCAAAATGCCAAAGTAGCTGATGATTTGACGCTTCCTAAATGAGGTAGAT 1260  
Db 1320 GGGGAGTCTGAATCAAAATGCCAAAGTAGCTGATGATTTGACGCTTCCTAAATGAGGTAGAT 1379  
Qy 1261 GAATATCTGCTGTTCTTCCAGAAAGTGAATGATGATGTTAGTGTGATGACTGCATGAT 1320  
Db 1380 GAATATCTGCTGTTCTTCCAGAAAGTGAATGATGATGTTAGTGTGATGACTGCATGAT 1439  
Qy 1321 ATATGTAAGTAAGTAAGAGTTCACCTCCAAATCAGTAGAGATTAATTTGAAGACAAATA 1380  
Db 1440 ATATGTAAGTAAGTAAGAGTTCACCTCCAAATCAGTAGAGATTAATTTGAAGACAAATA 1499  
Qy 1381 TTTTGGAAAACTATTCGGAAGAAAGCAAGCTCCCAACTTAAGCCATGTAACCTGAAAT 1440  
Db 1500 TTTTGGAAAACTATTCGGAAGAAAGCAAGCTCCCAACTTAAGCCATGTAACCTGAAAT 1559  
Qy 1441 CTATATTAAGAGATTTGTTACTGAGCAACAGATTAATACAGAGCCTCCCTCACAAT 1500  
Db 1560 CTATATTAAGAGATTTGTTACTGAGCAACAGATTAATACAGAGCCTCCCTCACAAT 1619  
Qy 1501 AAATTAAGCGTAAAGAGACCTACATCAGGCTTCATCCTGAGATTTTATCAAGAAA 1560  
Db 1620 AAATTAAGCGTAAAGAGACCTACATCAGGCTTCATCCTGAGATTTTATCAAGAAA 1679  
Qy 1561 GCAGATTTGGCAGTTCAAAAGACTCCTGAAATGATTAATCAGGAGACTTAACCAACGAG 1620  
Db 1680 GCAGATTTGGCAGTTCAAAAGACTCCTGAAATGATTAATCAGGAGACTTAACCAACGAG 1739  
Qy 1621 CAGAGTGTCAAGTAGATTAATTAATAGTGTGATGATGATTAATCAAGAAAGTGAT 1680  
Db 1740 CAGAGTGTCAAGTAGATTAATTAATAGTGTGATGATGATTAATCAAGAAAGTGAT 1799  
Qy 1681 TCTATTCAGATGAGAAATTCCTAACCCATAGATCAGTCCGAAAGAAATGCTTTTC 1740  
Db 1800 TCTATTCAGATGAGAAATTCCTAACCCATAGATCAGTCCGAAAGAAATGCTTTTC 1859  
Qy 1741 AAAACGAAGCTGAACCTATTAAGCAGCATTAAGCAATATGAACTCGAATTAATATC 1800  
Db 1860 AAAACGAAGCTGAACCTATTAAGCAGCATTAAGCAATATGAACTCGAATTAATATC 1919  
Qy 1801 CACAATTCAAAGCACTTAAGAAATAGGCTGAGAGGAAAGTCTTACAGGCAATTT 1860  
Db 1920 CACAATTCAAAGCACTTAAGAAATAGGCTGAGAGGAAAGTCTTACAGGCAATTT 1979  
Qy 1861 CAGCGCTGAACCTAGTAGTAGTAATTAAGCCACCTAATTTAGTACTGATTTCAA 1920  
Db 1980 CAGCGCTGAACCTAGTAGTAGTAATTAAGCCACCTAATTTAGTACTGATTTCAA 2039  
Qy 1921 ATTGATAGTTGTTCTAGCAGTGAAGAGATTAAGAAAAAAGTACAACCAATGCCAGTC 1980  
Db 2040 ATTGATAGTTGTTCTAGCAGTGAAGAGATTAAGAAAAAAGTACAACCAATGCCAGTC 2099  
Qy 1981 AGGCACGCAAGAAACCTACACATCTATGGAAGGTAAAGAACTGCACTGAGCCAGAG 2040  
Db 2100 AGGCACGCAAGAAACCTACACATCTATGGAAGGTAAAGAACTGCACTGAGCCAGAG 2159

Qy 2041 AGTAACAAGCCAAATGAACACAGCAAGTAAAGACATGACATGATCTTCCAGAGCT 2100  
Db 2160 AGTAACAAGCCAAATGAACACAGCAAGTAAAGACATGACATGATCTTCCAGAGCT 2219  
Qy 2101 AAGTTAACAAATGACACCTGGTTCTTTTACTAGTGTTCAAATACAGTGAACCTTAAGAA 2160  
Db 2220 AAGTTAACAAATGACACCTGGTTCTTTTACTAGTGTTCAAATACAGTGAACCTTAAGAA 2279  
Qy 2161 TTTGTCAATCCTAGCTTTCCAGAGAAAGAAAGAAAGAACTAGAAACAGTTAAAGTG 2220  
Db 2280 TTTGTCAATCCTAGCTTTCCAGAGAAAGAAAGAAAGAAAGAACTAGAAACAGTTAAAGTG 2339  
Qy 2221 TCTATATATGCTGAGAGACCCCAAGATCTCATGTTAAGTGAAGAGGTTTCCAACT 2280  
Db 2340 TCTATATATGCTGAGAGACCCCAAGATCTCATGTTAAGTGAAGAGGTTTCCAACT 2399  
Qy 2281 GAAAGATCTGTAGAGATGACATATTTCATTTGTTAGTGTGATGATGATGATGATGAT 2340  
Db 2400 GAAAGATCTGTAGAGATGACATATTTCATTTGTTAGTGTGATGATGATGATGATGATGAT 2459  
Qy 2341 GAAAGATCTGTAGAGATGACATATTTCATTTGTTAGTGTGATGATGATGATGATGATGAT 2400  
Db 2460 GAAAGATCTGTAGAGATGACATATTTCATTTGTTAGTGTGATGATGATGATGATGATGAT 2519  
Qy 2401 TGTGTAGTCAAGTGTGACATTTGAAACCCCAAGGACATTAATTCATGTTGTTCCAA 2460  
Db 2520 TGTGTAGTCAAGTGTGACATTTGAAACCCCAAGGACATTAATTCATGTTGTTCCAA 2579  
Qy 2461 GATATATGAATATGACACAGAAAGGCTTAAGTATCATTTGGACATGAAGTTAACACAGT 2520  
Db 2580 GATATATGAATATGACACAGAAAGGCTTAAGTATCATTTGGACATGAAGTTAACACAGT 2639  
Qy 2521 CGGGAACCAAGCATAGAAATGGAAGAAAGTCACTTGTATGTCAGTATTTCCAAATACA 2580  
Db 2640 CGGGAACCAAGCATAGAAATGGAAGAAAGTCACTTGTATGTCAGTATTTCCAAATACA 2699  
Qy 2581 TTCAGGTTTCAAGGCCCGCATATTTGCTCTGTTTCAAAATCAGAGAAATGCGAAGAG 2640  
Db 2700 TTCAGGTTTCAAGGCCCGCATATTTGCTCTGTTTCAAAATCAGAGAAATGCGAAGAG 2739  
Qy 2641 GAATGTCAACATCTCTGCCCATCTGGGTCTTTAAAGAAACCAAGTCCAAAGTCACT 2700  
Db 2760 GAATGTCAACATCTCTGCCCATCTGGGTCTTTAAAGAAACCAAGTCCAAAGTCACT 2819  
Qy 2701 TTTGAATGTGAACAAAGAAAGAAATCAAGGAAAGATGATGATTAATCAACCTGTA 2760  
Db 2820 TTTGAATGTGAACAAAGAAAGAAATCAAGGAAAGATGATGATTAATCAACCTGTA 2879  
Qy 2761 CAGACATTAATATCACTGACAGCTTCTGTTGTTGTCAGAAAGTAAAGCCAGTTGAT 2820  
Db 2880 CAGACATTAATATCACTGACAGCTTCTGTTGTTGTCAGAAAGTAAAGCCAGTTGAT 2939  
Qy 2821 AATGCCAAATGTATATCAAGAGAGGCTAGGTTTGTATCATCTCACTTTCAGAGGC 2880  
Db 2940 AATGCCAAATGTATATCAAGAGAGGCTAGGTTTGTATCATCTCACTTTCAGAGGC 2999  
Qy 2881 AAGCAACTGACATCATTTCCAAATTAACATGAGACTTTTACAAACCATATGCTATA 2940  
Db 3000 AAGCAACTGACATCATTTCCAAATTAACATGAGACTTTTACAAACCATATGCTATA 3059  
Qy 2941 CCACACATTTTCCCATCAAGTCAATTTGTTAAACCTAAATGTAAGAAATCTCTGAG 3000  
Db 3060 CCACACATTTTCCCATCAAGTCAATTTGTTAAACCTAAATGTAAGAAATCTCTGAG 3119  
Qy 3001 GAAAACTTTGAGGAACATTAATGTACCTGAAAGAGAAATGGGAAATGAGAACATTTCA 3060  
Db 3120 GAAAACTTTGAGGAACATTAATGTACCTGAAAGAGAAATGGGAAATGAGAACATTTCA 3179  
Qy 3061 AGTACATGAGCACAATTTAGCCGTAATTAACATTTAGAGAAATGTTTAAAGAACCCAGC 3120  
Db 3180 AGTACATGAGCACAATTTAGCCGTAATTAACATTTAGAGAAATGTTTAAAGAACCCAGC 3239

QY 3121 TCAAGCAATATTAAAGTAGTTCACAGTAACTAATGAATGGGCTCCAGTATTATGA 3180  
| | | | |  
Db 3240 TCAAGCAATATTAAAGTAGTTCACAGTAACTAATGAATGGGCTCCAGTATTATGA 3299  
| | | | |  
QY 3181 ATAGGTTCAGTATGATAAATCAATTCAGAGAACTAGTGAAGAACAGAGGCCAAATG 3240  
| | | | |  
Db 3300 ATAGGTTCAGTATGATAAATCAATTCAGAGAACTAGTGAAGAACAGAGGCCAAATG 3359  
| | | | |  
QY 3241 AATGCTATCTAGATTAGGGGTTTGGCAACCTGAGTCTATTAACAAAGCTTCCGGA 3300  
| | | | |  
Db 3360 AATGCTATCTAGATTAGGGGTTTGGCAACCTGAGTCTATTAACAAAGCTTCCGGA 3419  
| | | | |  
QY 3301 ACTAATTTAGACATCTGAAATTAAGAACCAAGATATGAGAAGTAGTTGACACTGT 3360  
| | | | |  
Db 3420 ACTAATTTAGACATCTGAAATTAAGAACCAAGATATGAGAAGTAGTTGACACTGT 3479  
| | | | |  
QY 3361 AATPACGATTTCTCCATATCTGATTTGATTAAGTGAACAGCTATGGGAGTAGT 3420  
| | | | |  
Db 3480 AATPACGATTTCTCCATATCTGATTTGATTAAGTGAACAGCTATGGGAGTAGT 3539  
| | | | |  
QY 3421 CATGATCTCAGGTTTGTCTGAGACACCTGATGACCTGTTAGATGATGGTGAATTAAG 3480  
| | | | |  
Db 3540 CATGATCTCAGGTTTGTCTGAGACACCTGATGACCTGTTAGATGATGGTGAATTAAG 3599  
| | | | |  
QY 3481 GAAGATACTAGTTTGTCTGAAATGACATTTAAGGAAAGTTCTGCTGTTTACCAAAAGC 3540  
| | | | |  
Db 3600 GAAGATACTAGTTTGTCTGAAATGACATTTAAGGAAAGTTCTGCTGTTTACCAAAAGC 3659  
| | | | |  
QY 3541 GTCGAAAGAGAGAGTTCAGAGAGTCCAGGCTTCACGCTTACACATTTGGGCTCAG 3600  
| | | | |  
Db 3660 GTCGAAAGAGAGAGTTCAGAGAGTCCAGGCTTCACGCTTACACATTTGGGCTCAG 3719  
| | | | |  
QY 3601 GGTACCGAAGAGGGGCAAGAAATTAAGTCTCTGAGAGAGAACTATCTAGTAGAGAT 3660  
| | | | |  
Db 3720 GGTACCGAAGAGGGGCAAGAAATTAAGTCTCTGAGAGAGAACTATCTAGTAGAGAT 3779  
| | | | |  
QY 3661 GAAGAGCTTCCCTGCTTCAACACTTGTATTGTGAAGTAACAAATATACCTTCCAG 3720  
| | | | |  
Db 3780 GAAGAGCTTCCCTGCTTCAACACTTGTATTGTGAAGTAACAAATATACCTTCCAG 3839  
| | | | |  
QY 3721 TCTACTAGGATAGACCGTGTGTACGAGTGTCTGTAGAACAAGAGGAAATTA 3780  
| | | | |  
Db 3840 TCTACTAGGATAGACCGTGTGTGTACGAGTGTCTGTAGAACAAGAGGAAATTA 3899  
| | | | |  
QY 3781 TTATCATTTGAAGAAATAGCTTAATAGACTCCAGTAACAGGATATATTGGCAAGGCACT 3840  
| | | | |  
Db 3900 TTATCATTTGAAGAAATAGCTTAATAGACTCCAGTAACAGGATATATTGGCAAGGCACT 3959  
| | | | |  
QY 3841 CAGGAACATCACCTTAGTGAAGAAACAATGTTCTGCTAGCTGTTTCTTACAGTGC 3900  
| | | | |  
Db 3960 CAGGAACATCACCTTAGTGAAGAAACAATGTTCTGCTAGCTGTTTCTTACAGTGC 4019  
| | | | |  
QY 3901 AGTGAATTTGGAAGACTTGACTGCAATTAACAAACAGGAGATCCTTCTGATTGATTC 3960  
| | | | |  
Db 4020 AGTGAATTTGGAAGACTTGACTGCAATTAACAAACAGGAGATCCTTCTGATTGATTC 4079  
| | | | |  
QY 3961 TCCAAACAATGAGGATCAGTCTGAAAGCCAGGGAGTTGCTGAGTGAACAAGAAATG 4020  
| | | | |  
Db 4080 TCCAAACAATGAGGATCAGTCTGAAAGCCAGGGAGTTGCTGAGTGAACAAGAAATG 4139  
| | | | |  
QY 4021 GTTTAGATGATGAAGAAAGAGAGAGGCTTGGAAAGAAATATCAAGAGACAAAGC 4080  
| | | | |  
Db 4140 GTTTAGATGATGAAGAAAGAGAGAGGCTTGGAAAGAAATATCAAGAGACAAAGC 4199  
| | | | |  
QY 4081 ATGATTTCAACTTATGAGTGAAGCAGCATGGGTGAGAGTGAACAAGCGCTCTGAA 4140  
| | | | |  
Db 4200 ATGATTTCAACTTATGAGTGAAGCAGCATGGGTGAGAGTGAACAAGCGCTCTGAA 4259  
| | | | |  
QY 4141 GACTGCTCAGGGCTATCTCTGAGAGTGAATTTTAACCACTCAGAGAGAGGATACCATG 4200  
| | | | |  
Db 4260 GACTGCTCAGGGCTATCTCTGAGAGTGAATTTTAACCACTCAGAGAGAGGATACCATG 4319  
| | | | |  
QY 4201 CAACATTAACCTGATAAGCTCCAGCAGAGAAATGGCTGAACCTGTGTTAGAACAG 4260  
| | | | |

Db 4320 CAACATTAACCTGATAAGCTCCAGCAGAAATGGCTGAACCTAGAGAGCTGTGTAGAACAG 4379  
| | | | |  
QY 4261 CATGGAGCGAGCTCTTAACAGAGTACCTTCATCATATAGTACCTTCCCTTACAG 4320  
| | | | |  
Db 4380 CATGGAGCGAGCTCTTAACAGAGTACCTTCATCATATAGTACCTTCCCTTACAG 4439  
| | | | |  
QY 4321 GACCTGCAAAATCCAGAACAAAGCACATCAGAAAAAGCAGTATTAATCTTCAGAAAAAGT 4380  
| | | | |  
Db 4440 GACCTGCAAAATCCAGAACAAAGCACATCAGAAAAAGCAGTATTAATCTTCAGAAAAAGT 4499  
| | | | |  
QY 4381 AGTGAATTAACCTTAAGCCAGAAATCCAGAGGCTTCTGCTGACAAAGTTGAGTGTCT 4440  
| | | | |  
Db 4500 AGTGAATTAACCTTAAGCCAGAAATCCAGAGGCTTCTGCTGACAAAGTTGAGTGTCT 4559  
| | | | |  
QY 4441 GCAGATAGTCTACCGATTAAGAAATTAAGAACAGGAGTGAAGAGGATCCCTCTPAA 4500  
| | | | |  
Db 4560 GCAGATAGTCTACCGATTAAGAAATTAAGAACAGGAGTGAAGAGGATCCCTCTPAA 4619  
| | | | |  
QY 4501 TGCCCATCATTTAGATGATAGTGTGTACATGACACAGTTGCTTGGAGTCTTCAGAAATGA 4560  
| | | | |  
Db 4620 TGCCCATCATTTAGATGATAGTGTGTACATGACACAGTTGCTTGGAGTCTTCAGAAATGA 4679  
| | | | |  
QY 4561 AACTACCATCTCAGAGAGAGCTCATTTAAGTTGTTGATGTGAGAGACACAGCTGAA 4620  
| | | | |  
Db 4680 AACTACCATCTCAGAGAGAGCTCATTTAAGTTGTTGATGTGAGAGACACAGCTGAA 4739  
| | | | |  
QY 4621 GAGTCTGGGCGACACGATTTGAAGGAAACATCTTCTGCAAGGCAAGATCTAGAGGA 4680  
| | | | |  
Db 4740 GAGTCTGGGCGACACGATTTGAAGGAAACATCTTCTGCAAGGCAAGATCTAGAGGA 4799  
| | | | |  
QY 4681 ACCCTTACCTGGAATTCGGAATTCAGCCCTTCTCTGATGACCCGTAATCTGATCCTCT 4740  
| | | | |  
Db 4800 ACCCTTACCTGGAATTCGGAATTCAGCCCTTCTCTGATGACCCGTAATCTGATCCTCT 4859  
| | | | |  
QY 4741 GAAGACAGAGCCCGAGAGTACGCTGTGTGGCAACATACCTTCTCAACCTCTGCAATG 4800  
| | | | |  
Db 4860 GAAGACAGAGCCCGAGAGTACGCTGTGTGGCAACATACCTTCTCAACCTCTGCAATG 4919  
| | | | |  
QY 4801 AAGTTTCCCAATTTGAAGTTGAGAAATCTGGCCAGAGTCCAGTGTGCTCATCTACT 4860  
| | | | |  
Db 4920 AAGTTTCCCAATTTGAAGTTGAGAAATCTGGCCAGAGTCCAGTGTGCTCATCTACT 4979  
| | | | |  
QY 4861 GATACGTCTGGTATTAATGCAATGGAAGAAAGTGTGAGCAGAGGAGACACGAAATGACA 4920  
| | | | |  
Db 4980 GATACGTCTGGTATTAATGCAATGGAAGAAAGTGTGAGCAGAGGAGACACGAAATGACA 5039  
| | | | |  
QY 4921 GCTTCAACGAAAGGTCACAAAGAAATGTCCATGTGTGTCTGGCTGACCCGAGAA 4980  
| | | | |  
Db 5040 GCTTCAACGAAAGGTCACAAAGAAATGTCCATGTGTGTCTGGCTGACCCGAGAA 5099  
| | | | |  
QY 4981 GAATTTATCTCTGTGACAAAGTTTGCAGAAACACACATCATCTTAACTAATTAAT 5040  
| | | | |  
Db 5100 GAATTTATCTCTGTGACAAAGTTTGCAGAAACACACATCATCTTAACTAATTAAT 5159  
| | | | |  
QY 5041 ACTGAAGACATACATCATGTTGTAAGAAACAGATGCTGAGTTTGTGTGAACGAGCA 5100  
| | | | |  
Db 5160 ACTGAAGACATACATCATGTTGTAAGAAACAGATGCTGAGTTTGTGTGAACGAGCA 5219  
| | | | |  
QY 5101 CTGAATATTTTCTAGGAATTTGGGAGAGAAATGGGTAGTTAGTATTTCTGGGTGACC 5160  
| | | | |  
Db 5220 CTGAATATTTTCTAGGAATTTGGGAGAGAAATGGGTAGTTAGTATTTCTGGGTGACC 5279  
| | | | |  
QY 5161 CAGTCTATTAAGAAAGAAATGCTGSAATGAGATGATTTTGAAGTCAAGAGGAGATG 5220  
| | | | |  
Db 5280 CAGTCTATTAAGAAAGAAATGCTGSAATGAGATGATTTTGAAGTCAAGAGGAGATG 5339  
| | | | |  
QY 5221 GTCAATGGAAGAACCCAGAGTCCAAAGCAGCAGAAAGATCCAGAGCAGAAAGATC 5280  
| | | | |  
Db 5340 GTCAATGGAAGAACCCAGAGTCCAAAGCAGCAGAAAGATCCAGAGCAGAAAGATC 5399  
| | | | |  
QY 5281 TTCAGGGGCTTGAATCTGTTGCTATGGGCTTCCACCAATGCTCCACAGATCACTG 5340  
| | | | |

Db 5400 TTCAGGGGGCTAGAAATCTGTGCTATGGGCCCTTCACCAACATGCCACAGATCAACTG 5459  
 Qy 5441 GAATGATGTGTAGACGTGTGTGCTTGTGTGTGAGAGAGCTTTCATTCACCTT 5400  
 Db 5460 GAATGATGTGTAGACGTGTGTGCTTGTGTGTGAGAGAGCTTTCATTCACCTT 5519  
 Qy 5401 GGCACAGGTGTCCACCCAAATTTGTGTGTGTGAGAGAGCTTTCATTCACCTT 5460  
 Db 5520 GGCACAGGTGTCCACCCAAATTTGTGTGTGTGAGAGAGCTTTCATTCACCTT 5579  
 Qy 5461 TTCCATGCAATTTGGGCGAGATGTGTGAGGACCTTGTGTGACCCGAGAGTGTGTGAG 5520  
 Db 5580 TTCCATGCAATTTGGGCGAGATGTGTGAGGACCTTGTGTGACCCGAGAGTGTGTGAG 5639  
 Qy 5521 AGTGTAGACTCTACCAATGCGGAGAGCTGTGAGGACCTTGTGTGACCCGAGAGTGTGTGAG 5580  
 Db 5640 AGTGTAGACTCTACCAATGCGGAGAGCTGTGAGGACCTTGTGTGACCCGAGAGTGTGTGAG 5699  
 Qy 5581 AGCCACTAC 5589  
 Db 5700 AGCCACTAC 5708

RESULT 12  
 176943 176943 5914 bp DNA Linear PAT 03-APR-1998  
 LOCUS Sequence 1 from patent US 5693473.  
 ACCESSION 176943  
 VERSION 176943.1 GI:3013097  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 5914)  
 AUTHORS Shattuck-Eidens,D.M., Simard,J., Durocher,F., Eml,M. and Nakamura,Y.  
 TITLE Linked breast and ovarian cancer susceptibility gene  
 JOURNAL Patent: US 5693473-A 1 02-DEC-1997;  
 FEATURES location/Qualifiers  
 source 1..5914  
 BASE COUNT 2006 a 1156 c 1316 g 1436 t  
 ORIGIN

Query Match 99.9%; Score 5585.8; DB 6; Length 5914;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 5587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 Qy 1 ATGATTTATCTGCTCTTCGCGTTGAGAGATGACAAATGTCATTTATGCTATGACAGAA 60  
 Db 120 ATGATTTATCTGCTCTTCGCGTTGAGAGATGACAAATGTCATTTATGCTATGACAGAA 179  
 Qy 61 ATCTTGAAGTGTCCCATCTGTCTGTGAGATGATCAAGAGACCTGTCTCCAAAGTGTGAC 120  
 Db 180 ATCTTGAAGTGTCCCATCTGTCTGTGAGATGATCAAGAGACCTGTCTCCAAAGTGTGAC 239  
 Qy 121 CACATATTTTGCAGATTTTGCATGCTGAACCTTCTCAACCAAGAGAGAGGCTTTCACAG 180  
 Db 240 CACATATTTTGCAGATTTTGCATGCTGAACCTTCTCAACCAAGAGAGAGGCTTTCACAG 299  
 Qy 181 TGTCTTATATGATGATATTAACCAAGAGAGCTTACAGAAATAGATGATGATTTAGT 240  
 Db 300 TGTCTTATATGATGATATTAACCAAGAGAGCTTACAGAAATAGATGATGATTTAGT 359  
 Qy 241 CAATCTTGAAGAGCTATTGAAGATCTTTGTCTTTCAGCTTGCACAGAGTTTGAG 300  
 Db 360 CAATCTTGAAGAGCTATTGAAGATCTTTGTCTTTCAGCTTGCACAGAGTTTGAG 419  
 Qy 301 TATGCAAAAGCTATATTTTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 360  
 Db 420 TATGCAAAAGCTATATTTTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 479  
 Qy 361 GAATTTCTATCATCCAAAGATATGGGCTACAGAAACCGTCCCAAGAGCTTCTACAGAT 420

Db 480 GAATTTCTATCATCCAAAGATATGGGCTACAGAAACCGTCCCAAGAGCTTCTACAGAT 539  
 Qy 421 GAACCCGAAATCTTCTCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 480  
 Db 540 GAACCCGAAATCTTCTCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 599  
 Qy 481 ACTGTGAGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 540  
 Db 600 ACTGTGAGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 659  
 Qy 541 GAATTTGGATCTGATTTCTTGTGAGAGATATCCGTTAATTAAGCACTTATTCAGTGTGGA 600  
 Db 660 GAATTTGGATCTGATTTCTTGTGAGAGATATCCGTTAATTAAGCACTTATTCAGTGTGGA 719  
 Qy 601 GATCAACAATTTGTTCAATTCACCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAT 660  
 Db 720 GATCAACAATTTGTTCAATTCACCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAT 779  
 Qy 661 GCAGAGAGAGCTGCTGTGTGAGATTTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAT 720  
 Db 780 GCAGAGAGAGCTGCTGTGTGAGATTTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAT 839  
 Qy 721 CCCAGTATATGATTTTGAACACACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 780  
 Db 840 CCCAGTATATGATTTTGAACACACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 899  
 Qy 781 TATCAGGGATGTTCTGTTTCAAACTGATGTGAGAGAGAGAGAGAGAGAGAGAGAT 840  
 Db 900 TATCAGGGATGTTCTGTTTCAAACTGATGTGAGAGAGAGAGAGAGAGAGAGAGAT 959  
 Qy 841 AGCTCATTAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 900  
 Db 960 AGCTCATTAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1019  
 Qy 901 AAGCTGAATTTCTGATTAATAAGCAACACCTGCTTACCAAGAGAGAGAGAGAGAT 960  
 Db 1020 AAGCTGAATTTCTGATTAATAAGCAACACCTGCTTACCAAGAGAGAGAGAGAGAT 1079  
 Qy 961 TGGGCTGGAAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1020  
 Db 1080 TGGGCTGGAAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1139  
 Qy 1021 GATCTGAATGCTGATCCCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1080  
 Db 1140 GATCTGAATGCTGATCCCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1199  
 Qy 1081 TCAGAGAGATCTAGAGATGATGAGAGATGCTTGTGATTAACACTAATATAGAGATTCAG 1140  
 Db 1200 TCAGAGAGATCTAGAGATGATGAGAGATGCTTGTGATTAACACTAATATAGAGATTCAG 1259  
 Qy 1141 AAGCTTAATGAGTGTTCCTCAGAGAGATGATGATGATGATGATGATGATGATGAT 1200  
 Db 1260 AAGCTTAATGAGTGTTCCTCAGAGAGATGATGATGATGATGATGATGATGATGAT 1319  
 Qy 1201 GGGGAGTGAATCAAAATGCAAGATGATGATGATGATGATGATGATGATGATGAT 1260  
 Db 1320 GGGGAGTGAATCAAAATGCAAGATGATGATGATGATGATGATGATGATGATGAT 1379  
 Qy 1261 GAATATTTCTGTTCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1320  
 Db 1380 GAATATTTCTGTTCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1439  
 Qy 1321 ATATGTAAGAGTGAAGAGTTCACCTCAATCACTAGAGAGATATTTAGAGAGAGAT 1380  
 Db 1440 ATATGTAAGAGTGAAGAGTTCACCTCAATCACTAGAGAGATATTTAGAGAGAGAT 1499  
 Qy 1381 TTTGGGAGAGAGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1440  
 Db 1500 TTTGGGAGAGAGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1559  
 Qy 1441 CTAAATTATAGAGACATTTGTTACTGAGAGAGAGAGAGAGAGAGAGAGAGAT 1500



Db	1560	CTAATTATAGACGATTTTGTACTGAGCCACAGATTAATACAGAGCGTCCCTCCACAAT	1619
QY	1501	AAATTAAAGCCGTAAGAGAGACCTCATCAGGCCCTTCATCCGTGAGGATTTTATCAGAAA	1560
Db	1620	AAATTAAAGCGTAAGAGAGAGACCTCATCAGGCCCTTCATCCGTGAGGATTTTATCAGAAA	1679
QY	1561	GCAGATTTGGCAGTTCAAAAGACTCCTGATAATGATTAATCAGGGAACTTAACCAACGGAG	1620
Db	1680	GCAGATTTGGCAGTTCAAAAGACTCCTGATAATGATTAATCAGGGAACTTAACCAACGGAG	1739
QY	1621	CAGATGGTCMAAGTATGAAATTTACTAATAGTGGTCTATGAGAAATTAACAAAGGTGAT	1680
Db	1740	CAGAAATGGTCMAAGTATGAAATTTACTAATAGTGGTCTATGAGAAATTAACAAAGGTGAT	1799
QY	1681	TCTATTCCAGAAATGAGAAAAATCCTTAACCCATAGAAATACATCGAAAAAAGATCTGCTTC	1740
Db	1800	TCTATTCCAGAAATGAGAAAAATCCTTAACCCATAGAAATACATCGAAAAAAGATCTGCTTC	1859
QY	1741	AAAAAGAAAGCTGAACCTATTAAGCAGCATTAAGCAATATGAAATCGAAATTAATATC	1800
Db	1860	AAAAAGAAAGCTGAACCTATTAAGCAGCATTAAGCAATATGAAATCGAAATTAATATC	1919
QY	1801	CACATTTCAAAAAGCAGCTTAAGAAAGATAGGCGTAGAGAGAAAGTCTTCAACAGCATATT	1860
Db	1920	CACATTTCAAAAAGCAGCTTAAAGAAATAGGCGTAGAGAGAAAGTCTTCAACAGCATATT	1979
QY	1861	CATGCGCTTGAACCTATAGTACGTAGAAATCTAAGCCCACTTAATTGACTGAATTGGCA	1920
Db	1980	CATGCGCTTGAACCTATAGTACGTAGAAATCTAAGCCCACTTAATTGACTGAATTGGCA	2039
QY	1921	ATTGATTAATGTTCTTACAGCATGAGAGATTAAGAAAAAGTACACCAATGCCAGTC	1980
Db	2040	ATTGATTAATGTTCTTACAGCATGAGAGATTAAGAAAAAGTACACCAATGCCAGTC	2099
QY	1981	AGGCACACAGAAACCTCACTCACTATGAGAGSTAAAGAACTCGCACTGGAGCCAAAGAG	2040
Db	2100	AGGCACACAGAAACCTCACTCACTATGAGAGSTAAAGAACTCGCACTGGAGCCAAAGAG	2159
QY	2041	AGTAACAAGCCAAATGAACAGACAGATTAAGAAAGACATGACAGTACTTCCAGAGCTG	2100
Db	2160	AGTAACAAGCCAAATGAACAGACAGATTAAGAAAGACATGACAGTACTTCCAGAGCTG	2219
QY	2101	AAAGTTAACAAATGCACCTGCTGTTCTTTACTAAGTGTTCAAATACCAGTGAACCTTAAGAA	2160
Db	2220	AAAGTTAACAAATGCACCTGCTGTTCTTTACTAAGTGTTCAAATACCAGTGAACCTTAAGAA	2279
QY	2161	TTTGTCATCTCAGAGCTTCGAAAGAGAAACAAAGAAAGTAAAGCAAGCTTAAGAGTG	2220
Db	2280	TTTGTCATCTCAGAGCTTCGAAAGAGAAACAAAGAAAGTAAAGCAAGCTTAAGAGTG	2339
QY	2221	TCTAATAATAGCTGAGAACCCCAAGAGATCTCATGTTAAGTGGAGAAAGGTTTCCAACT	2280
Db	2340	TCTAATAATAGCTGAGAACCCCAAGAGATCTCATGTTAAGTGGAGAAAGGTTTCCAACT	2399
QY	2281	GAAAGATCTGTAGAGAGTACGATATTTCAATGTTACCTGGTACTGATTAAGCAGCTAG	2340
Db	2400	GAAAGATCTGTAGAGAGTACGATATTTCAATGTTACCTGGTACTGATTAAGCAGCTAG	2459
QY	2341	GAAAGTATCTGCTACTAGGAGGTTAGACATCTAGGGAAGGCAAAAACAGAAACCAATATAA	2400
Db	2460	GAAAGTATCTGCTACTAGGAGGTTAGACATCTAGGGAAGGCAAAAACAGAAACCAATATAA	2519
QY	2401	TGTTGTGAGTCAGTGTGACGATTTTGAAAAACCCCAAGGACATTAATTCATGTTGTCCAAA	2460
Db	2520	TGTTGTGAGTCAGTGTGACGATTTTGAAAAACCCCAAGGACATTAATTCATGTTGTCCAAA	2579
QY	2461	GATTAATAGAAATGACACAGAAAGGCTTTAAGTATCCATTGGGACATGAAGTTAACACAGT	2520
Db	2580	GATTAATAGAAATGACACAGAAAGGCTTTAAGTATCCATTGGGACATGAAGTTAACACAGT	2639
QY	2521	CGGGAACACAGCATAGAAATGGAAGAAAGTGAACCTTGATGCTCAGTATTTGGACAAATACA	2580
Db	2640	CGGGAACACAGCATAGAAATGGAAGAAAGTGAACCTTGATGCTCAGTATTTGGACAAATACA	2699

QY	2581	TTCAAGGTTTCAAAGCCGACGTCAATTTGCTGTGTTTCAATCCAGGAATATGCAGACAG	2640
Db	2700	TTCAAGGTTTCAAAGCCGACGTCAATTTGCTGTGTTTCAATCCAGGAATATGCAGAGAG	2759
QY	2641	GAATATGCAGACCTTCTGCGCCACCTGSGGTCCTTAAGAAACAAGTCCAAAGTCACT	2700
Db	2760	GAATATGCAGACCTTCTGCGCCACCTGSGGTCCTTAAGAAACAAGTCCAAAGTCACT	2819
QY	2701	TTTGAATGTGAACAAAGGAAGAAATCAAGGAAGAATGAGTCAATATTAAGCCTGTGA	2760
Db	2820	TTTGAATGTGAACAAAGGAAGAAATCAAGGAAGAATGAGTCAATATTAAGCCTGTGA	2879
QY	2761	CAGACAGTTAATATACACTGCAGAGCCTTCTGCTGTGTTGCTCAGAAAGATAAACCACTGAT	2820
Db	2880	CAGACAGTTAATATACACTGCAGAGCCTTCTGCTGTGTTGCTCAGAAAGATAAACCACTGAT	2939
QY	2821	AATGCCAATGTAGTATCAAGAGAGGCTCTAGTGTGTTCTCATCATCTCACTTCAGAGGC	2880
Db	2940	AATGCCAATGTAGTATCAAGAGAGGCTCTAGTGTGTTCTCATCATCTCACTTCAGAGGC	2999
QY	2881	AACGAACTGACCTATTACTTCAATTAACATGACCTTTTACAAACCCATTCGTATA	2940
Db	3000	AACGAACTGACCTATTACTTCAATTAACATGACCTTTTACAAACCCATTCGTATA	3059
QY	2941	CCACACATTTTCCCATCAAGTCATTTGTTAAATCAATATGTAAGAAATATCTGCTAGAG	3000
Db	3060	CCACACATTTTCCCATCAAGTCATTTGTTAAATCAATATGTAAGAAATATCTGCTAGAG	3119
QY	3001	GAAACCTTGAGAGACATTCATGTCACCTGGAAGAAATGGAATGAGACATTTCA	3060
Db	3120	GAAACCTTGAGAGACATTCATGTCACCTGGAAGAAATGGAATGAGACATTTCA	3179
QY	3061	AGTACAGTACACACATTAAGCCGTATPAACATTTAGAGAAATGTTTTAAAGAACCCAGC	3120
Db	3180	AGTACAGTACACACATTAAGCCGTATPAACATTTAGAGAAATGTTTTAAAGAACCCAGC	3239
QY	3121	TCACCAATATTATGAAGTAGGTTCCAGTACTAATGAAGTGGGCTCCAGTATTATGAA	3180
Db	3240	TCACCAATATTATGAAGTAGGTTCCAGTACTAATGAAGTGGGCTCCAGTATTATGAA	3299
QY	3181	ATAGGTTCCAGTGAAGAAACATTCACACACACAGTATGAACAGAGGCCCAAAATTTG	3240
Db	3300	ATAGGTTCCAGTGAAGAAACATTCACACACACAGTATGAACAGAGGCCCAAAATTTG	3359
QY	3241	AATGCTATGCTTAGATTTAGGGGTTTTGGACCTGAGGTCCTATPAACAAGTCCTCTCGGA	3300
Db	3360	AATGCTATGCTTAGATTTAGGGGTTTTGGACCTGAGGTCCTATPAACAAGTCCTCTCGGA	3419
QY	3301	AGTAATTTGAAGCATCTGGAATATAAAGACAGAAATATGAAGAATAGTTTCAACCTGTT	3360
Db	3420	AGTAATTTGAAGCATCTGGAATATAAAGACAGAAATATGAAGAATAGTTTCAACCTGTT	3479
QY	3361	AATACAGATTTCTCTCCATATCTGATTCAGTATTAAGAACAGCCATGSGGAAGTATG	3420
Db	3480	AATACAGATTTCTCTCCATATCTGATTCAGTATTAAGAACAGCCATGSGGAAGTATG	3539
QY	3421	CATGCATCTCAGGTTTGTCTGAGACACCTGATGACCTGTTAGATGATGTGAATAAAG	3480
Db	3540	CATGCATCTCAGGTTTGTCTGAGACACCTGATGACCTGTTAGATGATGTGAATAAAG	3599
QY	3481	GAACTACTACTTTTGTGGAATAATGACATTAAGAAAGTTCTGCTGTTTTAGCAAAAGC	3540
Db	3600	GAACTACTACTTTTGTGGAATAATGACATTAAGAAAGTTCTGCTGTTTTAGCAAAAGC	3659
QY	3541	GTCCAGAAAGAGAGCTTAGCAGAGAGTCCTAGCCCTTTCACCCCTATACATTTTGGCTCAG	3600
Db	3660	GTCCAGAAAGAGAGCTTAGCAGAGAGTCCTAGCCCTTTCACCCCTATACATTTTGGCTCAG	3719
QY	3601	GGTTACCGAAGAGGGGCCAAGAATAATAGAGTCTCTCAGAAAGACATTATCTAGTAGAGAT	3660
Db	3720	GGTTACCGAAGAGGGGCCAAGAATAATAGAGTCTCTCAGAAAGACATTATCTAGTAGAGAT	3779



QY	3661	GAAGAGCTTCCGCTTCCTCCAAACCTGTCTPTTTGGTAAAGAAACATATACCTTCACG	3720
Db	3780	GAAGAGCTTCCGCTTCCTCCAAACCTGTCTPTTTGGTAAAGAAACATATACCTTCACG	3839
QY	3721	TCTACTAGGAGATAGCACGCTTGCTACCGAGTGTCTGTCTAAGAACACAGAGAGAAATTTA	3780
Db	3840	TCTACTAGGAGATAGCACGCTTGCTACCGAGTGTCTGTCTAAGAACACAGAGAGAAATTTA	3899
QY	3781	TTATCATTTGAAGATAGCTTAAATGACTGCGATACCAAGGTAAATTGGCAAGGCAATCT	3840
Db	3900	TTATCATTTGAAGATAGCTTAAATGACTGCGATACCAAGGTAAATTGGCAAGGCAATCT	3959
QY	3841	CAGGAACATCATCCTTAGTGGAGGAAACAAATGTTCTGTGTAGCTGTTTCTTCACATGCG	3900
Db	3960	CAGGAACATCATCCTTAGTGGAGGAAACAAATGTTCTGTGTAGCTGTTTCTTCACATGCG	4019
QY	3901	AGTAATTTGGAAAGCTTGACTGCAAAATACAAACCCAGAGTCCCTTCTTGATGTGTTCT	3960
Db	4020	AGTAATTTGGAAAGCTTGACTGCAAAATACAAACCCAGAGTCCCTTCTTGATGTGTTCT	4079
QY	3961	TCCAAACAAATGAGGCAATCAGTCTGAAAGCCAGGAGTGGTGTCTGAGTGCACAAAGATTG	4020
Db	4080	TCCAAACAAATGAGGCAATCAGTCTGAAAGCCAGGAGTGGTGTCTGAGTGCACAAAGATTG	4139
QY	4021	GTTTCACATGATGTAAGAAAGAGAGCGGCTTGGAGAAATTAATCAAGAGCAAAAGC	4080
Db	4140	GTTTCACATGATGTAAGAAAGAGAGCGGCTTGGAGAAATTAATCAAGAGCAAAAGC	4199
QY	4081	ATGGAATTCAAACTATAGGTGAAGGACGACATCGGGGTGTAGAGTGAACAAAGGTCTCGAA	4140
Db	4200	ATGGAATTCAAACTATAGGTGAAGGACGACATCGGGGTGTAGAGTGAACAAAGGTCTCGAA	4259
QY	4141	GACTGCTCAGGGCTATCCTCTCAGAGAGACATTTTAAACCACTCGACGAGAGGATACCATG	4200
Db	4260	GACTGCTCAGGGCTATCCTCTCAGAGAGACATTTTAAACCACTCGACGAGAGGATACCATG	4319
QY	4201	CAACATTAACCTGATTAAGCTCCACAGAGAAATGGCTAATCTAGAAGCTGTGTAGAACAG	4260
Db	4320	CAACATTAACCTGATTAAGCTCCACAGAGAAATGGCTAATCTAGAAGCTGTGTAGAACAG	4379
QY	4261	CATGGAGAGCCAGCCTTTTAACAGGTACCCTCCATCAATAGTGAAGCTCTTGCGCCCTGAG	4320
Db	4380	CATGGAGAGCCAGCCTTTTAACAGGTACCCTCCATCAATAGTGAAGCTCTTGCGCCCTGAG	4439
QY	4321	GACCTGGCAATCCAGAACAAAGCACATCCAGAAACACAGATTAATTAACCTACAGAAAGT	4380
Db	4440	GACCTGGCAATCCAGAACAAAGCACATCCAGAAACACAGATTAATTAACCTACAGAAAGT	4499
QY	4381	AGTGAATTAACCTTAAGCCAGAAATCCAGAAAGGCTTTCTGCTGCACAAAGTTGAGGTCT	4440
Db	4500	AGTGAATTAACCTTAAGCCAGAAATCCAGAAAGGCTTTCTGCTGCACAAAGTTGAGGTCT	4559
QY	4441	GCAGATAGTCTTCCAGATTAATAAATAAGAACCCAGAGATGGAAAGGTATCCCTCTTAA	4500
Db	4560	GCAGATAGTCTTCCAGATTAATAAATAAGAACCCAGAGATGGAAAGGTATCCCTCTTAA	4619
QY	4501	TGCCATCATTTAGATGATAGGTGGTATACATCACAGTTGCTGTGGAGTCTTCAGAAATAGA	4560
Db	4620	TGCCATCATTTAGATGATAGGTGGTATACATCACAGTTGCTGTGGAGTCTTCAGAAATAGA	4679
QY	4561	AACTAACCCATCTCAAGAGAGACTCATTTAAGGTGTTATGTGAGAGCAACAGCTGAA	4620
Db	4680	AACTAACCCATCTCAAGAGAGACTCATTTAAGGTGTTATGTGAGAGCAACAGCTGAA	4739
QY	4621	GAGCTGAGGCGACACGATTTGAGGGAACATCTTACTTGCACAGGCAAGATCTTAGAGGA	4680
Db	4740	GAGCTGAGGCGACACGATTTGAGGGAACATCTTACTTGCACAGGCAAGATCTTAGAGGA	4799
QY	4681	ACCCCTTAACCTGGAATCTGGAATCAGGCTCTTCTGATGACCCCTGGAATCTGATCTTCT	4740
Db	4800	ACCCCTTAACCTGGAATCTGGAATCAGGCTCTTCTGATGACCCCTGGAATCTGATCTTCT	4859
QY	4741	GAAGACAGAGCCCCAGAGTCAAGTCTGCTGTGGCAACATACCATCTTCAACCTTGCAATG	4800

D	b	4860	GAAGACAGAGCCCCAGAGCTCAGCTCGTGTGGCAACATACCATTCTTAACCTCTGCATTG	4919
Q	y	4801	AAAGTTCCCAATTGAAAGTTGCAGAAATCTGGCCAGAGTCCAGCTGCTCATACTACT	4860
D	b	4920	AAATCTCCCAATTTGAAGAATTGCAGAAATCTCCAGAGTCCAGCTGCTGCATACTACT	4979
Q	y	4861	GATACCTGCTGGGTATATGCATGTGAAGAAAGTGTGAGCAGGAGGAAGCCAGATTGACA	4920
D	b	4980	GATATCTCTGGGTATATGCAATGTGAAGAAAGTGTGAGCAGGAGGAAGCCAGAAATTGACA	5039
Q	y	4921	GCTTCAACAGAAAGGCGTCACAAAGAAATGCCATGGTGTGCTGTGCTGACCCAGAA	4980
D	b	5040	GCTTCAACAGAAAGGCTCAACAAAGAAATGTCATGGTGTGCTGTGCTGACCCAGAA	5099
Q	y	4981	GAATTATATGCTGTGTACAGATTGTGCAGAAACCCACATCACTTTAACTTAATCTAAT	5040
D	b	5100	GAATTATATGCTGTGTACAGATTGTGCAGAAACCCACATCACTTTAACTTAATCTAAT	5159
Q	y	5041	ACTGACAGACACTCTCATGTTGTTATGAAAAACAGATCTCTGATGTTGTGTGTAACGACA	5100
D	b	5160	ACTGAAAGACACTCTCATGTTGTTATGAAAAACAGATCTCTGATGTTGTGTGTAACGACA	5219
Q	y	5101	CTGAAATATTTTCTAGGAATTTGGGGAGGAAATGGGTAGTTAGCTATTTCTGGGTAC	5160
D	b	5220	CTGAAATATTTTCTAGGAATTTGGGGAGGAAATGGGTAGTTAGCTATTTCTGGGTAC	5279
Q	y	5161	CAGTCTATTTAAAGAAAGAAATATGCTGAATGAGCATGATTTTGAAGTCAGAGAGATGTG	5220
D	b	5280	CAGTCTATTTAAAGAAAGAAATATGCTGAATGAGCATGATTTTGAAGTCAGAGAGATGTG	5339
Q	y	5221	GTCATATGGAAGAAACCCACCAAGTGTCCAAAGCAGACAGACAGAAATTTCCAGACAGAAATC	5280
D	b	5340	GTCATATGGAAGAAACCCACCAAGTGTCCAAAGCAGACAGACAGAAATTTCCAGACAGAAATC	5399
Q	y	5281	TTTCAGGGGGGTAGAAATCTGTGCTATGAGGCGCTTCACCAACATGCCACAGATCACTG	5340
D	b	5400	TTTCAGGGGGGTAGAAATCTGTGCTATGAGGCGCTTCACCAACATGCCACAGATCACTG	5459
Q	y	5341	GAATGATGCTACAGCTGTGTGTGCTGTGTGTGTGAAGAGACTTTTCATCATTCACCTT	5400
D	b	5460	GAATGATGCTACAGCTGTGTGTGCTGTGTGTGTGAAGAGACTTTTCATCATTCACCTT	5519
Q	y	5401	GGCAGAGGTGCCACCCCAATTTGTGTGTGTGTCAGCCAAATGCTGTGACAGAGACAAATGGC	5460
D	b	5520	GGCAGAGGTGCCACCCCAATTTGTGTGTGTGTGTCAGCCAAATGCTGTGACAGAGACAAATGGC	5579
Q	y	5461	TTTCATCATTAATTTGGGCGAGATGTGTGAGGACACTTGTGTGACCCGAGAGTGGAGTTTGAC	5520
D	b	5580	TTTCATCATTAATTTGGGCGAGATGTGTGAGGACACTTGTGTGACCCGAGAGTGGAGTTTGAC	5639
Q	y	5521	AGTGTAGCACTCTACAGTGTGCCAGAGACTGGACACTTACTGATATACCCAGATCCCCAC	5580
D	b	5640	AGTGTAGCACTCTACAGTGTGCCAGAGACTGGACACTTACTGATATACCCAGATCCCCAC	5699
Q	y	5581	AGCCACTAC 5589	
D	b	5700	AGCCACTAC 5708	

RESULT	13			
180938				
LOCUS	180938	5914 bp	DNA	Linear
DEFINITION	Sequence 1 from patent US 5709999.			PAT 10-JUN-1998
ACCESSION	180938			
VERSION	180938.1	GI:3209228		
KEYWORDS				
SOURCE	Unknown.			
ORGANISM	Unknown.			
	Unclassified.			
REFERENCE	1 (bases 1 to 5914)			
AUTHORS	Statluch-Eldens,D.M., Simard,J., Durocher,F., Eml,M. and Nakamura,Y.			

TITLE Linked breast and ovarian cancer susceptibility gene  
JOURNAL Patent: US 5709999-A 1 20-JAN-1998:  
FEATURES Location/Qualifiers  
Source 1. 5914  
/organism="unknown"  
BASE COUNT 2006 a 1156 c 1316 g 1436 t  
ORIGIN  
Query Match 99.9%; Score 5585.8; DB 6; Length 5914;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 5587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 ATGATTTATCTGCTCTTCGCGCTTGAAGAATGCTATTATGCTATGAGAAA 60  
DB 120 ATGATTTATCTGCTCTTCGCGCTTGAAGAATGCTATTATGCTATGAGAAA 179  
QY 61 ATCTAGAGTCCCATCTGCTGAGTTCAGCAAGAACCTCTCCCAAAAGTGTAC 120  
DB 180 ATCTAGAGTCCCATCTGCTGAGTTCAGCAAGAACCTCTCCCAAAAGTGTAC 239  
QY 121 CACATATTTTGCATAATTTTGCATGCTGAACCTTCTCAACGAGAAAGGGCTTCACAG 180  
DB 240 CACATATTTTGCATAATTTTGCATGCTGAACCTTCTCAACGAGAAAGGGCTTCACAG 299  
QY 181 TGTCTTTATGTAGAATGATATTAACCAAAAGAGCCTACAGAAAGTACGAGATTAGT 240  
DB 300 TGTCTTTATGTAGAATGATATTAACCAAAAGAGCCTACAGAAAGTACGAGATTAGT 359  
QY 241 CAACCTGTGAGAGCTATGTAATTCATTGTGCTTTCACCTTACACAGGTTTGGAG 300  
DB 360 CAACCTGTGAGAGCTATGTAATTCATTGTGCTTTCACCTTACACAGGTTTGGAG 419  
QY 301 TATGCAAAAGCTATATATTTTGCATAAAAGAAATTAATCTCTGTAACATCTAAAGAT 360  
DB 420 TATGCAAAAGCTATATATTTTGCATAAAAGAAATTAATCTCTGTAACATCTAAAGAT 479  
QY 361 GAAGTTTCTATCATCCAAAGTATGGGCTACAGAAACCGTGCCTCAAAAGCTTCTACAGAT 420  
DB 480 GAAGTTTCTATCATCCAAAGTATGGGCTACAGAAACCGTGCCTCAAAAGCTTCTACAGAT 539  
QY 421 GAACCCGAAATCCTTCCTGCGAGAAACCACTGCTCAGTCCCACTCTTAACCTTGA 480  
DB 540 GAACCCGAAATCCTTCCTGCGAGAAACCACTGCTCAGTCCCACTCTTAACCTTGA 599  
QY 481 ACTGTGAGACCTCTGAGGACAAAGCAGGATACAACTCAAAAGACGCTCTGTACATT 540  
DB 600 ACTGTGAGACCTCTGAGGACAAAGCAGGATACAACTCAAAAGACGCTCTGTACATT 659  
QY 541 GAATTTGGATCTGATTTCTTGAGATACCGTTAATTAAGGCACTTATTCAGTGTGGA 600  
DB 660 GAATTTGGATCTGATTTCTTGAGATACCGTTAATTAAGGCACTTATTCAGTGTGGA 719  
QY 601 GATCAAGAATGTTTCAATCAATCCCTCAAGGAAACGAGGATGAATCGTTTGATTCT 660  
DB 720 GATCAAGAATGTTTCAATCAATCCCTCAAGGAAACGAGGATGAATCGTTTGATTCT 779  
QY 661 GCAAAAAAGGCTGCTTGTGAATTTCTGAGAGGATGTAACTAATCTAATCATCA 720  
DB 780 GCAAAAAAGGCTGCTTGTGAATTTCTGAGAGGATGTAACTAATCTAATCATCA 839  
QY 721 CCCAGTAATATGATTTGAACACCACTGAGAAAGCTGCAAGTGTGAGAGGATCCAGAAAG 780  
DB 840 CCCAGTAATATGATTTGAACACCACTGAGAAAGCTGCAAGTGTGAGAGGATCCAGAAAG 899  
QY 781 TATAGAGGTAGTCTGTTCAACCTGATGTGAGGCAATGTGAGCAAAATCTCATGCG 840  
DB 900 TATAGAGGTAGTCTGTTCAACCTGATGTGAGGCAATGTGAGCAAAATCTCATGCG 959  
QY 841 AGCTATTACAGCATGAGACAGAGCTTATTTACTACATAAGAGAGATGAATGTAGAA 900  
DB 960 AGCTATTACAGCATGAGACAGAGCTTATTTACTACATAAGAGAGATGAATGTAGAA 1019  
QY 901 AAGGCTGAATTTCTGTATATTAAGCAACAGCCTGCTTGAAGAGAGCCAACTAACA 960

DB 1020 AAGGCTGAATTTCTGTATATAAAGCAACAGCCTGGCTTACAGAGAGGCCAACATTAACA 1079  
QY 961 TGGGCTGGAAGTATAGAAAACATGTATGATAGCGGACTCCAGACAGAAAAAGGTA 1020  
DB 1080 TGGGCTGGAAGTATAGAAAACATGTATGATAGCGGACTCCAGACAGAAAAAGGTA 1139  
QY 1021 GATCTGAATCTGATCCCTGTGTGAGAGAAAAGATGAATTAACAGAACTGCAATGC 1080  
DB 1140 GATCTGAATCTGATCCCTGTGTGAGAGAAAAGATGAATTAACAGAACTGCAATGC 1199  
QY 1081 TCAGAGAACTCTAGAGATTAAGATGTTCTTGGATTAAGCTAATTAAGCAGATTAC 1140  
DB 1200 TCAGAGAACTCTAGAGATTAAGATGTTCTTGGATTAAGCTAATTAAGCAGATTAC 1259  
QY 1141 AAAGTTAATGAGTGGTTTCCAGAGAGTATGACAGTGTAGTTGCTGATGACTCAGAT 1200  
DB 1260 AAAGTTAATGAGTGGTTTCCAGAGAGTATGACAGTGTAGTTGCTGATGACTCAGAT 1319  
QY 1201 GGGAGTCTGAATCAAAATGCCAAAGTACGTATGATTGAGAGTTCTTAATGAGTATGAT 1260  
DB 1320 GGGAGTCTGAATCAAAATGCCAAAGTACGTATGATTGAGAGTTCTTAATGAGTATGAT 1379  
QY 1261 GAATATTTCTGTTCTTCAGAGAAATAGACTTACTGGCCAGTGTATCTCATGAGGCTTTA 1320  
DB 1380 GAATATTTCTGTTCTTCAGAGAAATAGACTTACTGGCCAGTGTATCTCATGAGGCTTTA 1439  
QY 1321 ATATGTAAAGCTGAAGAGATTCACCTCCAAATCAGTAGAGATTAATTTGAAGACAAATA 1380  
DB 1440 ATATGTAAAGCTGAAGAGATTCACCTCCAAATCAGTAGAGATTAATTTGAAGACAAATA 1499  
QY 1381 TTTGGAAACCTATACGAGACAGCAAGCCTCCCACTTAAGCATGTAACTGAATAAT 1440  
DB 1500 TTTGGAAACCTATACGAGAGAGCAAGCCTCCCACTTAAGCATGTAACTGAATAAT 1559  
QY 1441 CTAAATTAATAGAGCAATTTGTTACTGAGCCACAGATTAATTAAGAGCGTCCCTCAAAAT 1500  
DB 1560 CTAAATTAATAGAGCAATTTGTTACTGAGCCACAGATTAATTAAGAGCGTCCCTCAAAAT 1619  
QY 1501 AAATTAAGACCTTAAGAGAGACCTACATACAGCCCTTCCTGAGAGATTTATTAACAAA 1560  
DB 1620 AAATTAAGACCTTAAGAGAGACCTACATACAGCCCTTCCTGAGAGATTTATTAACAAA 1679  
QY 1561 GCAGATTTGGGAGTTCATAAAGACCTCTGAATATGATAATTCAGGGAACCTAACCAAGCGAG 1620  
DB 1680 GCAGATTTGGGAGTTCATAAAGACCTCTGAATATGATAATTCAGGGAACCTAACCAAGCGAG 1739  
QY 1621 CAGAATGCTCAAGTATGATTAATTAATAGTGTCAATGAGAAATTAACCAAAAGTGTAT 1680  
DB 1740 CAGAATGCTCAAGTATGATTAATTAATAGTGTCAATGAGAAATTAACCAAAAGTGTAT 1799  
QY 1681 TCTATTTCAAGATGAGAAAAATCCTAACCCATAGAAATCCTGAAAAAGATCTGCTTTC 1740  
DB 1800 TCTATTTCAAGATGAGAAAAATCCTAACCCATAGAAATCCTGAAAAAGATCTGCTTTC 1859  
QY 1741 AAAAGCAAAAGCTGAATCTATAAGCAGAGTATTAAGCAATATGAAATCTGAATTAATATC 1800  
DB 1860 AAAAGCAAAAGCTGAATCTATAAGCAGAGTATTAAGCAATATGAAATCTGAATTAATATC 1919  
QY 1801 CACAATTCAAAAGCACTTAAGAAAGATAGGCTGAGAGAGAGTCTTCTACAGGCAATATT 1860  
DB 1920 CACAATTCAAAAGCACTTAAGAAAGATAGGCTGAGAGAGAGTCTTCTACAGGCAATATT 1979  
QY 1861 CATGGCTTGAACCTGATGATGAGAAATTAAGCCACCTAATTTTACTGATGATTCGAA 1920  
DB 1980 CATGGCTTGAACCTGATGATGAGAAATTAAGCCACCTAATTTTACTGATGATTCGAA 2039  
QY 1921 ATTGATAGTCTTCTAGAGAGTGAAGATAAAGAAAAAAGTATCAACCAAAATGCCAGTCC 1980  
DB 2040 ATTGATAGTCTTCTAGAGAGTGAAGATAAAGAAAAAAGTATCAACCAAAATGCCAGTCC 2099  
QY 1981 AAGGACAGCAAAACCTAACAATCATGAGAGATTAAGAACTGTGCACTGAGCCAAAGAG 2040

OY	3121	TCACCAATATTATGAAGTAGTTCACGTAATAAGAGGGGCTCCAGTATTAAATGAA	3180
Db	3240	TCACCAATATTATGAAGTAGTTCACGTAATAAGAGGGGCTCCAGTATTAAATGAA	3299
OY	3181	ATAGTTCACGATGATGAACATTCACGACAGACTAGTACAAAACAGAGGCCAAAATTG	3240
Db	3300	ATAGTTCACGATGATGAACATTCACGACAGACTAGTACAAAACAGAGGCCAAAATTG	3359
OY	3241	AATGCTATGCTTACATTTAGGGGTTTTGCAACCTGAGGCTATAAACAAAGCTTCCTGGA	3300
Db	3360	AATGCTATGCTTACATTTAGGGGTTTTGCAACCTGAGGCTATAAACAAAGCTTCCTGGA	3419
OY	3301	AGTAATTTTGAAGCATTCCTGAATATAAAAGCAAGAAATGTGAAGAAGTAGTTCACAGCTGT	3360
Db	3420	AGTAATTTTGAAGCATTCCTGAATATAAAAGCAAGAAATGTGAAGAAGTAGTTCACAGCTGT	3479
OY	3361	AATACAGATTTCTCTCATATCTGATTTGAGATACTTAGAACAAGCCTATGGGAAGTAGT	3420
Db	3480	AATACAGATTTCTCTCATATCTGATTTGAGATACTTAGAACAAGCCTATGGGAAGTAGT	3539
OY	3421	CATGATCTCAGGTTTTGTTCTGAGACACCTATACCTGTTAGTGTATGGTAATAAG	3480
Db	3540	CATGATCTCAGGTTTTGTTCTGAGACACCTATACCTGTTAGTGTATGGTAATAAG	3599
OY	3481	GAGATCTACTATTTTGGCTGAATAATGACATTTAAGAAAGTTCTGCTGTTTTTACGAAAAGC	3540
Db	3600	GAGATCTACTATTTTGGCTGAATAATGACATTTAAGAAAGTTCTGCTGTTTTTACGAAAAGC	3659
OY	3541	GTCCAGAAAGSAGAGCTTAGCAGGAGTCTAGCCCTTTCACCCATPACATTTTGGCTAG	3600
Db	3660	GTCCAGAAAGSAGAGCTTAGCAGGAGTCTAGCCCTTTCACCCATPACATTTTGGCTAG	3719
OY	3601	GGTTACCGAAGAGGGGCCAACAAATTAGAGTCTCAGAGAGAACTTATCTAGTAGGAT	3660
Db	3720	GGTTACCGAAGAGGGGCCAACAAATTAGAGTCTCAGAGAGAACTTATCTAGTAGGAT	3779
OY	3661	GAGAGCTTCCCTGCTTCCACACCTGTTATTTGGTAAGTAACAAATATACCTTCCACAG	3720
Db	3780	GAGAGCTTCCCTGCTTCCACACCTGTTATTTGGTAAGTAACAAATATACCTTCCACAG	3839
OY	3721	TCTACTAGGCTAGACACGCTGTGCTACGAGTGTGTCTAAGAAACAGAGAGAAATTTA	3780
Db	3840	TCTACTAGGCTAGACACGCTGTGCTACGAGTGTGTCTAAGAAACAGAGAGAAATTTA	3899
OY	3781	TTATCATTTGAAGAAATAGCTTAATGACTGCACTAACCAAGGTAAATTTGGCAAGGCATCT	3840
Db	3900	TTATCATTTGAAGAAATAGCTTAATGACTGCACTAACCAAGGTAAATTTGGCAAGGCATCT	3959
OY	3841	CAGGAACATCCACTAGTAGAGAGAAACAAAGTCTGCTAGCTTTTCTTCACACTGC	3900
Db	3960	CAGGAACATCCACTAGTAGAGAGAAACAAAGTCTGCTAGCTTTTCTTCACACTGC	4019
OY	3901	AGTAATTTGAAGACTTGACTGCAAAATACAAACACCCAGAGTCCCTTCTTGATTTGGTCT	3960
Db	4020	AGTAATTTGAAGACTTGACTGCAAAATACAAACACCCAGAGTCCCTTCTTGATTTGGTCT	4079
OY	3961	TCCAAACAATGAGGCATCACTGTAAGAACCCAGGAGTTGGTCTGAGTGACACAGAAATTG	4020
Db	4080	TCCAAACAATGAGGCATCACTGTAAGAACCCAGGAGTTGGTCTGAGTGACACAGAAATTG	4139
OY	4021	GTTTCACAGTGTGAAGAAAGGAAAGCGGCTTGAACAAAATAATTCAAAGAACCAAGC	4080
Db	4140	GTTTCACAGTGTGAAGAAAGGAAAGCGGCTTGAACAAAATAATTCAAAGAACCAAGC	4199
OY	4081	ATGGAATTCAAACTTAGTGAAGCAGCATCTTGGTGAGTGAAGCAACAGCGTCTCGAA	4140
Db	4200	ATGGAATTCAAACTTAGTGAAGCAGCATCTTGGTGAGTGAAGCAACAGCGTCTCGAA	4259
OY	4141	GACTGCTCAGGGCTATCCTCTCAGAGTGACATTTTAAACACTCAGCAGAGGAGATACCATG	4200
Db	4260	GACTGCTCAGGGCTATCCTCTCAGAGTGACATTTTAAACACTCAGCAGAGGAGATACCATG	4319

4201 CAACATACTGATAAGCTCCAGCAGGAAATGGCTGAACAGTAAAGCTGTGTAGAACG 4260  
4320 CAACATACTGATAAGCTCCAGCAGGAAATGGCTGAACAGTAAAGCTGTGTAGAACG 4379  
4261 CATGGAGCCAGCCTTCTTAACAGCTACCTTCATCATTAAGTACTCTTCTGCCCTTGAG 4320  
4380 CATGGAGCCAGCCTTCTTAACAGCTACCTTCATCATTAAGTACTCTTCTGCCCTTGAG 4439  
4321 GACCTGGCAATCCAGAACAGACATCAGAAAAAGAGTATTAACCTTCACAGAAAGT 4380  
4440 GACCTGGCAATCCAGAACAGACATCAGAAAAAGAGTATTAACCTTCACAGAAAGT 4499  
4391 AGTGAATACCTTAAGCCAGAAATCCAGAAAGCCTTCTGCTGACAAAGTTTGAAGTCT 4440  
4500 AGTGAATACCTTAAGCCAGAAATCCAGAAAGCCTTCTGCTGACAAAGTTTGAAGTCT 4559  
4441 GCAGTACTTCTACCAAGTAAATTAAGAACAGAGTGAAGGATCATCCCTCTTAA 4500  
4560 GCAGTACTTCTACCAAGTAAATTAAGAACAGAGTGAAGGATCATCCCTCTTAA 4619  
4501 TGCCCATGATGATGATAGGTGTACATGCACAGTGTGCTGGGAGCTTCAGAAATGA 4560  
4620 TGCCCATGATGATGATAGGTGTACATGCACAGTGTGCTGGGAGCTTCAGAAATGA 4679  
4561 AACTACCATCTCAAGAGAGCTCATTAAGTTGTGATGTGAGAGACAGCCTGGA 4620  
4680 AACTACCATCTCAAGAGAGCTCATTAAGTTGTGATGTGAGAGACAGCCTGGA 4739  
4621 GAGTCTGGGCCACAGATTTGACGAAACATTTACTTCCAGAGCAGATCTAGAGGA 4680  
4740 GAGTCTGGGCCACAGATTTGACGAAACATTTACTTCCAGAGCAGATCTAGAGGA 4799  
4681 ACCCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4740  
4800 ACCCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4859  
4741 GAAGACAGAGCCCAAGCTCAGCTGTGTGCAACATACATCTTCAACCTCTGATG 4800  
4860 GAAGACAGAGCCCAAGCTCAGCTGTGTGCAACATACATCTTCAACCTCTGATG 4919  
4801 AAAGTCTCCCAATGAAGTTGCAAGTGTGCGGAGTCCAGCTGTGCTATCTACT 4860  
4920 AAAGTCTCCCAATGAAGTTGCAAGTGTGCGGAGTCCAGCTGTGCTATCTACT 4979  
4861 GATACGTGGGTATATGCAATGGAAGAAAGTGTGACAGAGGAGAGCAGAAATGGA 4920  
4980 GATACGTGGGTATATGCAATGGAAGAAAGTGTGACAGAGGAGAGCAGAAATGGA 5039  
4921 GCTTCAACAGAAAGGCTCAACAAAGAAATGTCATGTGTGTGCTGAGCCCTGAGAA 4980  
5040 GCTTCAACAGAAAGGCTCAACAAAGAAATGTCATGTGTGTGCTGAGCCCTGAGAA 5099  
4981 GAATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5040  
5100 GAATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5159  
5041 ACTGAAGAGTACTACTGTTGTTATGAAGACAGATGTGATTTGTGTGAAGGACA 5100  
5160 ACTGAAGAGTACTACTGTTGTTATGAAGACAGATGTGATTTGTGTGAAGGACA 5219  
5101 CTGAATATATTTCTAGGAATTCGCGGAGGAAATGGGTAGTATTTCTGGGTGACC 5160  
5220 CTGAATATATTTCTAGGAATTCGCGGAGGAAATGGGTAGTATTTCTGGGTGACC 5279  
5161 CAGCTATTTAAGAAAGAAATGCTGATGAGCATGATTTTGAAGTCAAGAGAGATGTG 5220  
5280 CAGCTATTTAAGAAAGAAATGCTGATGAGCATGATTTTGAAGTCAAGAGAGATGTG 5339  
5221 GTCAATGGAAGAAACCAAGAGTCCAAAGGAGAGAGAGATCCAGAGAGAAAGATC 5280  
5340 GTCAATGGAAGAAACCAAGAGTCCAAAGGAGAGAGAGATCCAGAGAGAAAGATC 5399  
5381 TTCAGGGGGCTAGAAATCTGTTGCTATGCGCTTTCACCAACATGCCCCAGATCAACTG 5340

5400 TTCAGGGGGCTAGAAATCTGTTGCTATGCGCCCTTACCAACATGCCAGATCAACTG 5459  
5341 GAATGAGTGTACACGTGTGTGCTCTTCTGTGTGAAGAGCTTTCATCATTCACCTT 5400  
5460 GAATGAGTGTACACGTGTGTGCTCTTCTGTGTGAAGAGCTTTCATCATTCACCTT 5519  
5401 GGCACAGGTGTACACCAATGTGTGTGACAGCAGATGCTGACAGAGACAAATGAGC 5460  
5520 GGCACAGGTGTACACCAATGTGTGTGACAGCAGATGCTGACAGAGACAAATGAGC 5579  
5461 TTCATGCAATTTGGCAGATGTGTGAGGACCTGTGTGACCCGAGAGTGGTGTGAC 5520  
5580 TTCATGCAATTTGGCAGATGTGTGAGGACCTGTGTGACCCGAGAGTGGTGTGAC 5639  
5521 AGTGTAGCACTTACCAGTCCAGAGCTGTGACACCTGATACCTGATACCCAGATCC 5580  
5640 AGTGTAGCACTTACCAGTCCAGAGCTGTGACACCTGATACCTGATACCCAGATCC 5699  
5581 AGCCACTAC 5589  
5700 AGCCACTAC 5708

RESULT 14  
LOCUS 181034 5914 bp DNA linear PAT 10-JUN-1998  
DEFINITION Sequence 1 from patent US 5710001.  
ACCESSION 181034  
VERSION 181034.1 GI:3209324  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 5914)  
AUTHORS Skolnick,M.H., Goldgar,D.E., Miki,Y., Swenson,J., Kamb,A., Hershman,K.D., Shattuck-Eidens,D.M., Tavtiglian,S.V., Wiseman,R.W., 17q-Linked breast and ovarian cancer susceptibility gene Patent: US 5710001-A 1 20-JAN-1998;  
FEATURES  
source Location/Qualifiers  
1..5914  
BASE COUNT. 2006 a 1156 c 1316 g 1436 t  
ORIGIN  
Query Match 99.9%; Score 5585.8; DB 6; Length 5914;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 5587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
1 ATGATTTATCTGCTCGGCTTGAGAGTACAAATGCAATTATGCTATGCAGAA 60  
120 ATGATTTATCTGCTCGGCTTGAGAGTACAAATGCTATGCTATGCAGAA 179  
61 ATCTTAGAGTGTCCATCTCTGAGTGTATCAAGAACCTGTCTCACAAGTGTGAC 120  
180 ATCTTAGAGTGTCCATCTCTGAGTGTATCAAGAACCTGTCTCACAAGTGTGAC 239  
121 CACATATTTTCAAAATTTTGCATGCTGAACCTCTCAACGAGAAAGGCTTCACAG 180  
240 CACATATTTTCAAAATTTTGCATGCTGAACCTCTCAACGAGAAAGGCTTCACAG 299  
181 TGCTCTTATGTAGATGATATACCAAAAGAGCTTCAAGAAAGTACGATTTAGT 240  
300 TGCTCTTATGTAGATGATATACCAAAAGAGGCTTCAAGAAAGTACGATTTAGT 359  
241 CAACCTTGTGAAGAGCTATTGAAATATCTTGTGCTTTAGTGTGACAGGTTGGAG 300  
360 CAACCTTGTGAAGAGCTATTGAAATATCTTGTGCTTTAGTGTGACAGGTTGGAG 419  
301 TATGCAAAACAGCTTAATTTTTCAGAAAAAGAAATTAACCTCTCTGACAACTTAAGAT 360  
420 TATGCAAAACAGCTTAATTTTTCAGAAAAAGAAATTAACCTCTCTGACAACTTAAGAT 479

OY	361	GAA GTT TCT ATC ATC CCA AAG TAT GGG C C TAC A GAA A C C G T G C C A A A A G A C T T C T A C A G A G T	420
Db	480	GAAG T T C T A T A C A T C C A A A G T A T G G C C T A C A G A A A C C C T G C C A A A A G A C T T C A C A G A G T	539
OY	421	GA C C C C G A A A A T C C T C C T T G C A G A A A C C A G T C A T G T C C A A C T C T A C C T T G S G A	480
Db	540	GA C C C C G A A A A T C C T T C C T T G C A G A A A C C A G T C T C A G T G T C C A C T C T A C A C C T T G S G A	599
OY	481	A C T G T G A M A C T C T G A G A C A A A G C A G C G G A T A C A A C C T C A A A A G A C C T G T C T A C A T T	540
Db	600	A C T G T G A M A C T C T G A G A C A A A G C A G C G G A T A C A A C C T C A A A A A G A C C T G T C T A C A T T	659
OY	541	GA A T T G G A T C G A T T C T T C T T G A A G A T A C C G T T A T A T A G C A C A C T T A T T G C A G T S G S G A	600
Db	660	GA A T T G G A T C G A T T C T T C T T G A A G A T A C C G T T A T A T A G C A A C T T A T T G G A G T G S G A	719
OY	601	G A T C A A G A A T T G T T A C A A A T A C A C C C C T C A A G A A A C C A G A G A T G A A N T A G T T T G G A T T C T	660
Db	720	G A T C A A G A A T T G T T A C A A A T A C A C C C C T C A A G A A A C C A G A G A T G A A N T A G T T T G G A T T C T	779
OY	661	G C A A A A A A A G C T G C T T G T G A A T T T T C T A G A C G A T G T A A C A A A T A C T G A A A T C A T C A A	720
Db	780	G C A A A A A A A G C T G C T T G T G A A T T T T C T A G A C G A T G T A A C A A A A T A C G A A A C A T C A T C A A	839
OY	721	C C A G T A T A T A T G A T T T G A M A C A C C A C T A G A A A G C T G A G C T G A G A G C A T C C A G A A A A G	780
Db	840	C C C A G T A T A T A T G A T T T G A M A C A C C A C A C T A G A A A G C T G C A G C T G A G A G C A T C C A G A A A A G	899
OY	781	T A T C A G G G T A G T C T G T T T C A A A C T T G C A T G T G A C C A T G T G C A C A A A T A C T A C A T C C	840
Db	900	T A T C A G G G T A G T C T G T T C A A A C T T G C A T G T G A C C A T G T G C A C A A A T A C T A C A T C C	959
OY	841	A G C T A T T A C A G C A N T G A A A C A G A C A G T T A T T A C T A C A T A A A G A C A G A T A T S T A A A	900
Db	960	A G C T A T T A C A G C A N T G A A A C A G A C A G T T A T T A C T A C A T A A A A G A C A G A T A A T G T A A A	1019
OY	901	A A G C T G A A T T C T G T A A T A A A A G C A A A C A G C T G C T T A G C A A G A G A G C C A A C A T A A C A G A	960
Db	1020	A A G C T G A A T T C T G T A A T A A A A A G C A A A C A G C T G C T T A G C A A G A G A G C C A A C A T A A C A G A	1079
OY	961	T G G C G T G A A T A A G A A A C A T G T A T A T A T A T A G C G A C T C C C A G C A C A G A A A A A A A G S T A	1020
Db	1080	T G G C G T G A A T A A G A A A C A T G T A T A T A T A T A G C G A C T C C C A G A C A C A G A A A A A A A G S T A	1139
OY	1021	G A T C T G A A T G C T G A T C C C C T G T G A G A A A A A G A A T G A N T A A G C A A A A C T G C A T G C	1080
Db	1140	G A T C T G A A T G C T G A T C C C C T G T G A G A A A A A G A A T G A N T A A G C A A A A C T G C A T G C	1199
OY	1081	T C A G A G A A T C T A G A G A T A C T G A A G A N T T C C T T G G A T A A C A C T A A A T A G C A G A C A T T C A G	1140
Db	1200	T C A G A G A A T C T A G A G A T A C T G A A G A N T T C C T T G G A T A A C A C T A A A T A G C A G A C A T T C A G	1259
OY	1141	A A A G T A A T A G T G T T T T C C A G A A G A T A A C A C T A A G T T C T G A T G A C A C T C A C A T G A T	1200
Db	1260	A A A G T A A T A G T G T T T T C C A G A A G A T A A C A C T A A G T T C T G A T G A C A C T C A C A T G A T	1319
OY	1201	G G G A G A G T C G A N T C A A A T C C C A A G T A G C T A T G T A T G A C G T T C T A A A T A G A G G T A G A T	1260
Db	1320	G G G A G A G T C G A N T C A A A T C C C A A G T A G C T A T G T A T G A C G T T C T A A A T A G A G G T A G A T	1379
OY	1261	G A A T A T T C T G T T C T T C A G A G A A A A T A G A C T T A C T G C C A G T A T C C T C A T A G A G C T T T T A	1320
Db	1380	G A A T A T T C T G T T C T T C A G A G A A A A T A G A C T T A C T G C C A G T A T C C T C A T A G A G C T T T T A	1439
OY	1321	A T A T G T A A A A T G A A A G A T C A C T C C C A A T C A G A G A G A T A A T A T T G A A G A C A A A A T A	1380
Db	1440	A T A T G T A A A A T G A A A G A G T C A C T C C C A A T C A G A G A G A T A A T A T T G A A G A C A A A A T A	1499
OY	1381	T T T G G G A A A A C C T A T C G G A G A A G C A A G C C T C C C A A C T T A A G C A T G T A A C T G A A A T	1440
Db	1500	T T T G G G A A A A C C T A T C G G A G A A G C A A G C C T C C C A A C T T A A G C A T G T A A C T G A A A T	1559

QY	1441	CTAATTATAGAGCACTTTGTTACTAGAGCCACAGATTAATACAGAGACGGTCCCTCACAAAT	1500
Db	1560	CTAAATTATAGAGCACTTTGTTACTAGAGCCACAGATTAATACAGAGACGGTCCCTCACAAAT	1619
QY	1501	AAATTAAGCCGTAAGAGAGACCTACATCAGAGCCCTCATCCTGAGATTTTATCAAGAA	1560
Db	1620	AAATTAAAGCCTAAAGAGAGACCTACATCAGAGCCCTCATCCTGAGATTTTATCAAGAA	1679
QY	1561	GCAGATTTGGCAGTTCCAAAGACTCTCGAATATGATTAATCAGGAGACTAACCAACGGAG	1620
Db	1680	GCAGATTTGGCAGTTCCAAAGACTCTCGAATATGATTAATCAGGAGACTAACCAACGGAG	1739
QY	1621	CAGAAATGCTCAAGTGAATGATTAATTAATTAATGCTCATGAGATTAACCAAAAGTGAT	1680
Db	1740	CAGAAATGCTCAAGTGAATGATTAATTAATTAATGCTCATGAGATTAACCAAAAGTGAT	1799
QY	1681	TCTATTTCGAATGAGAAAATATCTTAACCCATTAATATCATCTCGAATAAGAAATCGCTTC	1740
Db	1800	TCTATTTCGAATGAGAAAATATCTTAACCCATTAATATCATCTCGAATAAGAAATCGCTTC	1859
QY	1741	AAAAGCAAGCTGAACTTATTAAGCAGCAGTATTAAGCATATGAGAACTCGAATTAATATC	1800
Db	1860	AAAAGCAAGCTGAACTTATTAAGCAGCAGTATTAAGCATATGAGAACTCGAATTAATATC	1919
QY	1801	CACAATTTAAAAGCACCTAAAAGATATGCTGAGAGAGATCTTCTACCAGCCATATT	1860
Db	1920	CACAATTTAAAAGCACCTAAAAGATATGCTGAGAGAGATCTTCTACCAGCCATATT	1979
QY	1861	CATGGCGTTGAACATAGTATGATGAGAAATCTAAGCCACCTATTCTACGATTTGCCAA	1920
Db	1980	CATGGCGTTGAACATAGTATGATGAGAAATCTAAGCCACCTATTCTACGATTTGCCAA	2039
QY	1921	ATTGATATGTTTCTTACGAGTGAAGAGATTAAGAAAAAAAGTCAACCAATCCAGTC	1980
Db	2040	ATTGATATGTTTCTTACGAGTGAAGAGATTAAGAAAAAAAGTCAACCAATCCAGTC	2099
QY	1981	AGGCACAGCAGAAACCTTACAACATCATGGAAGGTAAAGAACCTGCACTGAGCCAAAG	2040
Db	2100	AGGCACAGCAGAAACCTTACAACATCATGGAAGGTAAAGAACCTGCACTGAGCCAAAG	2159
QY	2041	AGTAAACAAGCCAAATGAAACAGACAAAGTAAAGACATGACATGACTTTCCAGACTG	2100
Db	2160	AGTAAACAAGCCAAATGAAACAGACAAAGTAAAGACATGACATGACTTTCCAGACTG	2219
QY	2101	AAGTTAAACAATGCAACCTGCTTCTTAACTAATGTTCAATTAACACAGTAAAGAA	2160
Db	2220	AAGTTAAACAATGCAACCTGCTTCTTAACTAATGTTCAATTAACACAGTAAAGAA	2279
QY	2161	TTTGTCAATCCTAGCCCTTCCAAAGAGAAAGAAAGAAAGAAACTAGAAACAGTTAAAGTG	2220
Db	2280	TTTGTCAATCCTAGCCCTTCCAAAGAGAAAGAAAGAAAGAAACTAGAAACAGTTAAAGTG	2339
QY	2221	TCTAATTAATGCTGAAGACCCCAAGATCTCATGTTAAGTGAAGAAGGTTTTCGAACCT	2280
Db	2340	TCTAATTAATGCTGAAGACCCCAAGATCTCATGTTAAGTGAAGAAGGTTTTCGAACCT	2399
QY	2281	GAAAGATCTTAGAGAGTAGAGCATTAATTCATTTGTTACTCGTACTGATTTATGGCAGTCAG	2340
Db	2400	GAAAGATCTTAGAGAGTAGAGCATTAATTCATTTGTTACTCGTACTGATTTATGGCAGTCAG	2459
QY	2341	GAAAGTATCTCTGTTACTGGAAGTTAGACACTCTAGGAGAGGCAAAAACAGAACCAATATAA	2400
Db	2460	GAAAGTATCTCTGTTACTGGAAGTTAGACACTCTAGGAGAGGCAAAAACAGAACCAATATAA	2519
QY	2401	TGTGTGAGTCAAGTGTGACGACTTTTGAABAAACCCCAAGGACTAATTCAATGTTGTTCCAA	2460
Db	2520	TGTGTGAGTCAAGTGTGACGACTTTTGAABAAACCCCAAGGACTAATTCAATGTTGTTCCAA	2579
QY	2461	GATATATGAATGACACAGAAAGGCTTTTAAGTATCCATTTGGACACATGAAAGTAAACACAGT	2520
Db	2580	GATATATGAATGACACAGAAAGGCTTTTAAGTATCCATTTGGACACATGAAAGTAAACACAGT	2639
QY	2521	CGGGAACCAACCATAGAAATGGAAGAAAGTGAACCTTATGCTCATGTTTTCAGAAATACA	2580

```
|||||
Db 2640 CCGGAACAGCATAGAAATGGAAGAAAGGAACCTGATGCTCAGATATTGGCAGAAATACA 2699
OY 2581 TTCAAGGTTCAAAAGGCCAGTCATTTGCTCTTTTCAAAATCCAGGAATGCAGAAAG 2640
Db 2700 TTCAAGGTTCAAAAGGCCAGTCATTTGCTCTTTTCAAAATCCAGGAATGCAGAAAG 2759
OY 2641 GAATGTCAGCAATTCCTGCCACCTCTGGGTCTTAAAGAAACAAAGTCCAAAAGTCACT 2700
Db 2760 GAATGTCAGCAATTCCTGCCACCTCTGGGTCTTAAAGAAACAAAGTCCAAAAGTCACT 2819
OY 2701 TTTGAATGTGAACAAAAGAGAAATCAAGGAAGATGCTAATATCAAGCTCTGA 2760
Db 2820 TTTGAATGTGAACAAAAGAGAAATCAAGGAAGATGCTAATATCAAGCTCTGA 2879
OY 2761 CAGAGATTAATATCATCTGAGGCTTCTGCTGGTGGTCAGAAATATAGCCAGTTGAT 2820
Db 2880 CAGAGATTAATATCATCTGAGGCTTCTGCTGGTGGTCAGAAATATAGCCAGTTGAT 2939
OY 2821 AATGCCAAATGTAGTATCAAGAGAGGCTCTAGGTTTGTCTATCATCTCAGTTCCAGAGC 2880
Db 2940 AATGCCAAATGTAGTATCAAGAGAGGCTCTAGGTTTGTCTATCATCTCAGTTCCAGAGC 2999
OY 2881 AAGGAACTGGACTCTATCTCCAAATTAACATGACCTTTACAAACCCTATTCGTATA 2940
Db 3000 AAGGAACTGGACTCTATCTCCAAATTAACATGACCTTTACAAACCCTATTCGTATA 3059
OY 2941 CCAGCACTTTTCCATCAAGTCAATTTGTTAAACCTAATGTAGAAATCTGCTAGAG 3000
Db 3060 CCAGCACTTTTCCATCAAGTCAATTTGTTAAACCTAATGTAGAAATCTGCTAGAG 3119
OY 3001 GAAAACTTTGAGAGACATTCATGTCACCTGAAAGAAATGGGAATGGAACATTTCCA 3060
Db 3120 GAAAACTTTGAGAGACATTCATGTCACCTGAAAGAAATGGGAATGGAACATTTCCA 3179
OY 3061 AGTACAGTACAGCAATTTAGCCGTAAATACATTAGAGAAATTTTAAAGAGCCAGC 3120
Db 3180 AGTACAGTACAGCAATTTAGCCGTAAATACATTAGAGAAATTTTAAAGAGCCAGC 3239
OY 3121 TCAGCAATATTAATGAAGTAGTTCCAGTACTAATGAAGTGGCTCCAGTATTATTAAG 3180
Db 3240 TCAGCAATATTAATGAAGTAGTTCCAGTACTAATGAAGTGGCTCCAGTATTATTAAG 3299
OY 3181 ATAGGTTCCAGTATGAAAAACATTCAGAGAGAACTAGAGAAACAGAGGCCCAAAATTC 3240
Db 3300 ATAGGTTCCAGTATGAAAAACATTCAGAGAGAACTAGAGAAACAGAGGCCCAAAATTC 3359
OY 3241 AATGCTATGCTTAGATTTAGGGTTTTCGAACCTGAGGCTCTATTAACAAAGTCTTCTGGA 3300
Db 3360 AATGCTATGCTTAGATTTAGGGTTTTCGAACCTGAGGCTCTATTAACAAAGTCTTCTGGA 3419
OY 3301 AGTATTTGAAGCATCTGAAATTAAGAAAGCAAGATATGAAGATAGTTCGACTGTT 3360
Db 3420 AGTATTTGAAGCATCTGAAATTAAGAAAGCAAGATATGAAGATAGTTCGACTGTT 3479
OY 3361 AATACAGATTTCTCCATATCTGATTTAGATTAAGTATGAAACAGCCCTATGGAGTAGT 3420
Db 3480 AATACAGATTTCTCCATATCTGATTTAGATTAAGTATGAAACAGCCCTATGGAGTAGT 3539
OY 3421 CATGATATCAGGTTTGTCTGAGACACCTGATGACTGTTAGATATGTTGAATTAAG 3480
Db 3540 CATGATATCAGGTTTGTCTGAGACACCTGATGACTGTTAGATATGTTGAATTAAG 3599
OY 3481 GAAGATAGTATTTGCTGAAGATGCAATTAAGAAAGTTCGCTTTTGAAGCAAG 3540
Db 3600 GAAGATAGTATTTGCTGAAGATGCAATTAAGAAAGTTCGCTTTTGAAGCAAG 3659
OY 3541 GTCCAGAAAGAGAGGTTAGCAGAGAGTCTTAGCCCTTTCACCATACATTTGGCTCAG 3600
Db 3660 GTCCAGAAAGAGAGGTTAGCAGAGAGTCTTAGCCCTTTCACCATACATTTGGCTCAG 3719
OY 3601 GGTTCACCAAGAGGGGCCAAGAAATTAGAGTCTCGAAGAGAACTTATCTAGTAGAT 3660
|||||
Db 3720 GGTTCACCAAGAGGGGCCAAGAAATTAGAGTCTCGAAGAGAACTTATCTAGTAGAT 3779
OY 3661 GAAGAGCTTCCCTGCTTCCAAACCTGTATTTGGAAAGTAAACAAATATCTTCAG 3720
Db 3780 GAAGAGCTTCCCTGCTTCCAAACCTGTATTTGGAAAGTAAACAAATATCTTCAG 3839
OY 3721 TCTACTAGGATATGACACCGCTTGCTACCGAGTCTCTGTAAAGAACACAGAGAGAAATTA 3780
Db 3840 TCTACTAGGATATGACACCGCTTGCTACCGAGTCTCTGTAAAGAACACAGAGAGAAATTA 3899
OY 3781 TTATCATTTGAAGATATGCTTAATATGACTGACTAATACAGGTAATTTGGCAAGCATCT 3840
Db 3900 TTATCATTTGAAGATATGCTTAATATGACTGACTAATACAGGTAATTTGGCAAGCATCT 3959
OY 3841 CAGGAACATCACCTTGTGAGGAAACAAATGTGCTGAGCTGTTTCTTCACAGTGC 3900
Db 3960 CAGGAACATCACCTTGTGAGGAAACAAATGTGCTGAGCTGTTTCTTCACAGTGC 4019
OY 3901 AGTGAATTTGGAAGACTTTGACTGCAAAATACAAACACCCAGAGATCTTTCTGATTTGCT 3960
Db 4020 AGTGAATTTGGAAGACTTTGACTGCAAAATACAAACACCCAGAGATCTTTCTGATTTGCT 4079
OY 3961 TCCAAACAAATGAGGATCAGTCTGAAGCCAGGAGTTGGTCTAGTGACAGAAATTC 4020
Db 4080 TCCAAACAAATGAGGATCAGTCTGAAGCCAGGAGTTGGTCTAGTGACAGAAATTC 4139
OY 4021 GTTTCAGATGATGAAGAAAGAGAGAACGGGCTTGGAAAGAAATATCAAGAGAGCAAGC 4080
Db 4140 GTTTCAGATGATGAAGAAAGAGAGAACGGGCTTGGAAAGAAATATCAAGAGAGCAAGC 4199
OY 4081 ATGATATTCAACTTAGGTAGAGCAGATCTGGGTGTGAGAGTGAAGAAACAGGCTCTGAA 4140
Db 4200 ATGATATTCAACTTAGGTAGAGCAGATCTGGGTGTGAGAGTGAAGAAACAGGCTCTGAA 4259
OY 4141 GACTGCTCAGGGGTATCTCTCTCAGAGTGACATTTTAACCACTCAGCAGAGGATTCCTG 4200
Db 4260 GACTGCTCAGGGGTATCTCTCTCAGAGTGACATTTTAACCACTCAGCAGAGGATTCCTG 4319
OY 4201 CAACATTAACCTGATTAAGCTCCAGCAGAGAAATGGCTGAACCTGAGAGCTGTGTAAGAC 4260
Db 4320 CAACATTAACCTGATTAAGCTCCAGCAGAGAAATGGCTGAACCTGAGAGCTGTGTAAGAC 4379
OY 4261 CATGGAGCCAGGCTTCTAACAGCTAACCTTCCATCATTAAGTACTTTCGCTTGCCTGAG 4320
Db 4380 CATGGAGCCAGGCTTCTAACAGCTAACCTTCCATCATTAAGTACTTTCGCTTGCCTGAG 4439
OY 4321 GACCTGCGAAATCCAGAAAGACATCAGAAAGAGATTTAAGTTCACAGAAAGT 4380
Db 4440 GACCTGCGAAATCCAGAAAGACATCAGAAAGAGATTTAAGTTCACAGAAAGT 4499
OY 4381 AGTGAATACCTTATTAAGCCAGATCCAGAAAGGCTTTCGCTGACAACTTTGAGTGTCT 4440
Db 4500 AGTGAATACCTTATTAAGCCAGATCCAGAAAGGCTTTCGCTGACAACTTTGAGTGTCT 4559
OY 4441 GCAGATAGTCTTACACGTAATAAATAAGAACAGAGTGGAAAGTCAATCCCTTTTAA 4500
Db 4560 GCAGATAGTCTTACACGTAATAAATAAGAACAGAGTGGAAAGTCAATCCCTTTTAA 4619
OY 4501 TGCCCATCTTATGATGATGATGAGTGAACATGACATGCTCTGGAGTCTTCAAAATGA 4560
Db 4620 TGCCCATCTTATGATGATGATGAGTGAACATGACATGCTCTGGAGTCTTCAAAATGA 4679
OY 4561 AACTACCATCTCAAGAGAGGCTCTTAAAGTGTGTAAGTGGAGAGACACAGCTGGAA 4620
Db 4680 AACTACCATCTCTCAAGAGAGGCTCTTAAAGTGTGTAAGTGGAGAGACACAGCTGGAA 4739
OY 4621 GAGTCTGGGCCACAGATTTGACGGAACATCTTACTTGCCAAAGCAAGATCTAGAGGA 4680
Db 4740 GAGTCTGGGCCACAGATTTGACGGAACATCTTACTTGCCAAAGCAAGATCTAGAGGA 4799
OY 4681 ACCCTTACCTGGAATCTGGAATTCAGCCCTTCTGTGATGACCTGAATCTGATCTTCT 4740
Db 4800 ACCCTTACCTGGAATCTGGAATTCAGCCCTTCTGTGATGACCTGAATCTGATCTTCT 4859
|||||
```



OY		4741	GAAACAGAGGCCCCAGAGTGCAGCTCGTGTTGGCAACAATACCATCTTCACCTCTGCATGTG	4800
Dd		4860	GAAACAGAGGCCCCAGAGTGCAGCTCGTGTTGGCAACAATACCATCTTCACCTCTGCATGTG	4919
OY		4801	AAAGTTCOCCCATAATTGAAGAATTGCAGAAATCTGCCAGAGTCCAGCTGCCTGCATACTACT	4860
Dd		4920	AAAGTTCOCCCATAATTGAAGAATTGCAGAAATCTGCCAGAGTCCAGCTGCCTGCATACTACT	4979
OY		4881	GATACCTGCTGGGTATATATGCAATGTGAAAGAAAGTGTGAGCAGGAGGAGAACCCAGAAATTGACA	4920
Dd		4980	GATACCTGCTGGGTATATATGCAATGTGAAAGAAAGTGTGAGCAGGAGGAGAACCCAGAAATTGACA	5039
OY		4921	GCTTCAACAGAAAAGGCTCAACAAAAAAGATGTGCCATGCTGGTGTCTGGCCTACCCAGAA	4988
Dd		5040	GCTTCAACAGAAAAGGCTCAACAAAAAAGATGTGCCATGCTGGTGTCTGGCCTACCCAGAA	5099
OY		4981	GAATTTATGCTCGTGTAACAAGTTTGGCCAGAAAACACCACATCACCTTTACTAAATCTAATT	5048
Dd		5100	GAATTTATGCTCGTGTAACAAGTTTGGCCAGAAAACACCACATCACCTTTACTAAATCTAATT	5155
OY		5041	ACTGAGAGACTACTCATGTTGTTATGAAAAACAGATGCTGAGTTGTGTGTAACGGACA	5100
Dd		5160	ACTGAGAGAGACTACTCATGTTGTTATGAAAAACAGATGCTGAGTTGTGTGTAACGGACA	5219
OY		5101	CTGAAATTTTTTCTAGGAATTTGCCGGAGAGAAATTTGGTAGTTAGTACTATTTCTGGGTACC	5160
Dd		5220	CTGAAATTTTTTCTAGGAATTTGCCGGAGAGAAATTTGGTAGTTAGTACTATTTCTGGGTACC	5279
OY		5161	CAGCTATTTAAAGAAAGAAATCTGTAATGAGCATGATTTTGAAGTCAGAGAGAGATGTG	5220
Dd		5280	CAGCTATTTAAAGAAAGAAATCTGTAATGAGCATGATTTTGAAGTCAGAGAGAGATGTG	5339
OY		5221	GTCATATGAGAGAAACCCAAAGGTCCAAAGCGAGCAAGAGATCCCAAGACAGAAAGATC	5280
Dd		5340	GTCATATGAGAGAAACCCAAAGGTCCAAAGCGAGCAAGAGATCCCAAGACAGAAAGATC	5399
OY		5281	TTCAGGGGGCGTAGAAATCTGTTGCTATGAGGGCCCTTCAACCAATGCCACAGATCACTGTG	5340
Dd		5400	TTCAGGGGGCGTAGAAATCTGTTGCTATGAGGGCCCTTCAACCAATGCCACAGATCACTGTG	5459
OY		5341	GAATGAGATGTGACAGCTGTGTGTGCTCTCTGTGTTGAAGAGAGCTTTCATCTACCCCTT	5400
Dd		5460	GAATGAGATGTGACAGCTGTGTGTGCTCTCTGTGTTGAAGAGAGCTTTCATCTACCCCTT	5519
OY		5401	GGCACAGGTGTCCACCCAATTTGTGTTGTGTCAGCCAGATGCTCTGAGACAGAGCAATGGC	5460
Dd		5520	GGCACAGGTGTCCACCCAATTTGTGTTGTGTCAGCCAGATGCTCTGAGACAGAGCAATGGC	5579
OY		5461	TTCCATGCAATTTGGGACAGATGTGTGAGGACACTGTGGTGGAGCCCGAGATGGGTCTGTGAC	5520
Dd		5580	TTCCATGCAATTTGGGACAGATGTGTGAGGACACTGTGGTGGAGCCCGAGATGGGTCTGTGAC	5639
OY		5521	AGTGTAGACACTCTACAGTGTGCCAGAGAGCTGAGACACTACCTGATACCCAGATCCCCAC	5580
Dd		5640	AGTGTAGACACTCTACAGTGTGCCAGAGAGCTGAGACACTACCTGATACCCAGATCCCCAC	5699
OY		5581	AGCCACTAC 5589	
Dd		5700	AGCCACTAC 5708	
<b>RESULT_15</b>				
LOCUS	AR048660	5711 bp	DNA	linear PAT 29-SEP-1999
DEFINITION	Sequence 4 from patent US 5821328.			
ACCESSION	AR048660			
VERSION	AR048660.1 GI:5971003			
KEYWORDS				
SOURCE	Unknown.			
ORGANISM	Unknown.			
REFERENCE	1 (bases 1 to 5711)			

Query Match	Best Local Similarity	99.9% ; Score 5584.2 ; DB 6 ; Length 5711 ;	Matches 5586 ;	Conservative	0 ; Mismatches	3 ; Indels	0 ; Gaps	0 ;
1	ATGATATTTATTCGCTCTTCGCGTTGAGAGATACAAATGCTATTAATCTATGACAAA	60						
120	ATGATATTTATTCGCTCTTCGCGTTGAGAGATACAAATGCTATTAATCTATGACAAA	179						
61	ATCTTAGAGTGTCCATCTGTCTGAGTTGATCAAGAACTGTCTCCACAAAGTGTAC	120						
180	ATCTTAGAGTGTCCATCTGTCTGAGTTGATCAAGAACTGTCTCCACAAAGTGTAC	239						
121	CACATATTTTGCATATTTGATGCTGAACCTTCACACGAAAGAGGCGCTTCACAG	180						
240	CACATATTTTGCATATTTGATGCTGAACCTTCACACGAAAGAGGCGCTTCACAG	299						
181	TGTCCTTATGATAGATGATATTAACCAAGAGAGCCCTACAGAAAGTACGAGATTAGT	240						
300	GGTCTTTATGTAGAAATGATATTAACCAAGAGAGCCCTACAGAAAGTACGAGATTAGT	359						
241	CAACTTGTGGAAGCTATTTGAAAATCATTTGTGCTTTTACGCTTGACACAGGTTTGAG	300						
360	CAACTTGTGGAAGCTATTTGAAAATCATTTGTGCTTTTACGCTTGACACAGGTTTGAG	419						
301	TATGCAAAACACTATATTTTGGCAAAAAAGAAATTAATCTCTCTGAAATCTTAAAGAT	360						
420	TATGCAAAACACTATATTTTGGCAAAAAAGAAATTAATCTCTCTGAAATCTTAAAGAT	479						
361	GAGTTTATATCAATCAAGATGAGGCTACAGAAAACCTGCGCAAAAGACTTCTACAGAT	420						
480	GAGTTTATATCAATCAAGATGAGGCTACAGAAAACCTGCGCAAAAGACTTCTACAGAT	539						
421	GAACCCGAAAATCCTTCCTTCAGAGAAAACAGCTCAGTGTCCAACTCTTACACCTTGA	480						
540	GAACCCGAAAATCCTTCCTTCAGAGAAAACAGCTCAGTGTCCAACTCTTACACCTTGA	599						
481	ACTGTGAGAACTCTGAGAGCAAAAGCAGCGGATACAACTCTCAAAGAGCTGTCTACTT	540						
600	ACTGTGAGAACTCTGAGAGCAAAAGCAGCGGATACAACTCTCAAAGAGCTGTCTACTT	659						
541	GAAATGGGATCTGATTCCTTCTGGAAGATACGCTTATTAAGCACTTATTTGAGTGTGGA	600						
660	GAAATGGGATCTGATTCCTTCTGGAAGATACGCTTATTAAGCACTTATTTGAGTGTGGA	719						
601	GATCAGAAATGTTTACAAATCACCCCTCAAGAACACAGGATGGAATCACTTTGGATTCT	660						
720	GATCAGAAATGTTTACAAATCACCCCTCAAGAACACAGGATGGAATCACTTTGGATTCT	779						
661	GCAAAAAAGCGCTGCTGTGGAATTTTCTGAGACGATGTATACAAATTAAGCAATCATCAA	720						
780	GCAAAAAAGCGCTGCTGTGGAATTTTCTGAGACGATGTATACAAATTAAGCAATCATCAA	839						
721	CCCGATATATATGATTTGAAACACACAGAGAGGCTGAGTGTGAGGATCCAGAAAG	780						
840	CCCGATATATATGATTTGAAACACACAGAGAGGCTGAGTGTGAGGATCCAGAAAG	899						
781	TATAGAGGTAGTCTGTTTCAAACTTGCATGTGAGGCATGTGACACAAATATCATGCC	840						
900	TATAGAGGTAGTCTGTTTCAAACTTGCATGTGAGGCATGTGACACAAATATCATGCC	959						
841	AGCTCATTTAAGCATGAGAACAGCAGTTTATTAATCTCTAAAGACAGAAATGATGAAA	900						
960	AGCTCATTTAAGCATGAGAACAGCAGTTTATTAATCTCTCTAAAGACAGAAATGATGAAA	1019						



QY 901 AAGCTGATTCCTGTAATTAAGCAAAACAGCCTGGCTTAGCAAGGAGCCACATATACAGA 960  
DB 1020 AAGCTGATTCCTGTAATTAAGCAAAACAGCCTGGCTTAGCAAGGAGCCACATATACAGA 1079  
QY 961 TGGCTGGAAGTAAAGAAACATGTAAATGATAGCGGACTCCAGCACAGAAAAAAGTAA 1020  
DB 1080 TGGCTGGAAGTAAAGAAACATGTAAATGATAGCGGACTCCAGCACAGAAAAAAGTAA 1139  
QY 1021 GATCGAATGCGATCCCTGTGTAGAGAAAAAGATGAAGAAAGTAAAGTAAAGTAAAGTAA 1080  
DB 1140 GATCGAATGCGATCCCTGTGTAGAGAAAAAGATGAAGAAAGTAAAGTAAAGTAAAGTAA 1199  
QY 1081 TCAGAGAAATCCAGAGATACAGAGATGTTCTTGATTAACACTAAATAGCAGCATTCAG 1140  
DB 1200 TCAGAGAAATCCAGAGATACAGAGATGTTCTTGATTAACACTAAATAGCAGCATTCAG 1259  
QY 1141 AAAGTAAATGAGTGGTTTTCCAGAAAGTGAATGTTAGTTGATGATGATGATGATGAT 1200  
DB 1260 AAAGTAAATGAGTGGTTTTCCAGAAAGTGAATGTTAGTTGATGATGATGATGATGAT 1319  
QY 1201 GGGGAGTCTGATCAATGCGCAAGTGAAGTGAATGATGATGATGATGATGATGATGAT 1260  
DB 1320 GGGGAGTCTGATCAATGCGCAAGTGAAGTGAATGATGATGATGATGATGATGATGAT 1379  
QY 1261 GAATATTCCTGTTTTCTTCAGAGAAAAATGACTTACTGCGCAGTATCTCATGAGGCTTAA 1320  
DB 1380 GAATATTCCTGTTTTCTTCAGAGAAAAATGACTTACTGCGCAGTATCTCATGAGGCTTAA 1439  
QY 1321 ATATGTAAGTGAAGAGATTCCTCCAAATAGTAAATGATGATGATGATGATGATGAT 1380  
DB 1440 ATATGTAAGTGAAGAGATTCCTCCAAATAGTAAATGATGATGATGATGATGATGAT 1499  
QY 1381 TTTGGGAAAGCCTATCGAGAGAGCAAGCCTCCCAACTTAAGCATGATGATGATGAT 1440  
DB 1500 TTTGGGAAAGCCTATCGAGAGAGCAAGCCTCCCAACTTAAGCATGATGATGATGATGAT 1559  
QY 1441 CTAAATATAGAGATTTTGTACTGAGCCACAGATTAATCAAGAGCCTCCCTCAAAAT 1500  
DB 1560 CTAAATATAGAGATTTTGTACTGAGCCACAGATTAATCAAGAGCCTCCCTCAAAAT 1619  
QY 1501 AAATTAAGCCTTAAGAGAGAGCCTATCAGAGCCTTCACTCCAGAGATTTTATCAAGAA 1560  
DB 1620 AAATTAAGCCTTAAGAGAGAGCCTATCAGAGCCTTCACTCCAGAGATTTTATCAAGAA 1679  
QY 1561 GCAGATTTGGCAGTTCAAAAGAGCTCCTGAATGATTAATCAGGAGTAAACCAAAAGGAG 1620  
DB 1680 GCAGATTTGGCAGTTCAAAAGAGCTCCTGAATGATTAATCAGGAGTAAACCAAAAGGAG 1739  
QY 1621 CAGAATGCTCAAGTGAATTAATTAATAGTGTGATGAGAAATTAACCAAAAGGAGTAT 1680  
DB 1740 CAGAATGCTCAAGTGAATTAATTAATAGTGTGATGAGAAATTAACCAAAAGGAGTAT 1799  
QY 1681 TCTATTCAGATGAGAAAAATCCTTAACCAATAGATCAGTGGAAAAAGATCTGCTTTC 1740  
DB 1800 TCTATTCAGATGAGAAAAATCCTTAACCAATAGATCAGTGGAAAAAGATCTGCTTTC 1859  
QY 1741 AAAAGCAAGCTGAACCTTAAGAGAGAGTAAAGCAATATGGAAGTGAATTAATATATC 1800  
DB 1860 AAAAGCAAGCTGAACCTTAAGAGAGAGTAAAGCAATATGGAAGTGAATTAATATATC 1919  
QY 1801 CACAATTCMAAAGACCTTAAGAGAGAGTAAAGCAATATGGAAGTGAATTAATATATC 1860  
DB 1920 CACAATTCMAAAGACCTTAAGAGAGAGTAAAGCAATATGGAAGTGAATTAATATATC 1979  
QY 1861 CATCGCTGTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 1920  
DB 1980 CATCGCTGTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 2039  
QY 1921 ATTGATAGTGTCTAGCAGTGAAGAGATTAAGAAAAAAGTAAACCAATGCGCAGTC 1980  
DB 2040 ATTGATAGTGTCTAGCAGTGAAGAGATTAAGAAAAAAGTAAACCAATGCGCAGTC 2099

QY 1981 AGGCACAGCAGAAACCTTACACTCATGGAAGTAAAGAACTGCACTGAGCCAAAG 2040  
DB 2100 AGGCACAGCAGAAACCTTACACTCATGGAAGTAAAGAACTGCACTGAGCCAAAG 2159  
QY 2041 AGTAACAGCCCAATATGAACAGCAAGTAAAGACATGACAGTATCTTCCAGAGCTG 2100  
DB 2160 AGTAACAGCCCAATATGAACAGCAAGTAAAGACATGACAGTATCTTCCAGAGCTG 2219  
QY 2101 AAGTTAAACAAATGCACTGCTGCTTCTTACTAAGTGTCAATATACAGCACTTAAGAA 2160  
DB 2220 AAGTTAAACAAATGCACTGCTGCTTCTTACTAAGTGTCAATATACAGCACTTAAGAA 2279  
QY 2161 TTTGTCAATCCTAGCCTTCCAAAGAGAAAAAGAGAACTAGAAACAGTTAAAGTG 2220  
DB 2280 TTTGTCAATCCTAGCCTTCCAAAGAGAAAAAGAGAACTAGAAACAGTTAAAGTG 2339  
QY 2221 TCTAATTAATGCTGAAGACCCCAAGATCTCATGTTAACTGAGAAAGGTTTTGCAACT 2280  
DB 2340 TCTAATTAATGCTGAAGACCCCAAGATCTCATGTTAACTGAGAAAGGTTTTGCAACT 2399  
QY 2281 GAAAGATCTGTAGAGAGTAAAGCAGTATTTCAATTTGCTACTGATTAATGCGACTAG 2340  
DB 2400 GAAAGATCTGTAGAGAGTAAAGCAGTATTTCAATTTGCTACTGATTAATGCGACTAG 2459  
QY 2341 GAAAGATCTGTAGAGAGTAAAGCAGTATTTCAATTTGCTACTGATTAATGCGACTAG 2400  
DB 2460 GAAAGATCTGTAGAGAGTAAAGCAGTATTTCAATTTGCTACTGATTAATGCGACTAG 2519  
QY 2401 TGTGTGAGTCACTGTGCAAGCATTTTGAAGAACCCCAAGGACTAATTCATGTTTCCAA 2460  
DB 2520 TGTGTGAGTCACTGTGCAAGCATTTTGAAGAACCCCAAGGACTAATTCATGTTTCCAA 2579  
QY 2461 GATTAATGAAGTGAACAGAGAGGCTTAAAGTATGCAATGGAAGTAAACCAAGT 2520  
DB 2580 GATTAATGAAGTGAACAGAGAGGCTTAAAGTATGCAATGGAAGTAAACCAAGT 2639  
QY 2521 CGGGAACAGCATGAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 2580  
DB 2640 CGGGAACAGCATGAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 2699  
QY 2581 TTTCAAGTTTTCAAAAGGCGCAGTATTTGCTGTTTCAATCCAGGAATGCAAGAG 2640  
DB 2700 TTTCAAGTTTTCAAAAGGCGCAGTATTTGCTGTTTCAATCCAGGAATGCAAGAG 2759  
QY 2641 GAATGTGCAACATTTCTGCCCAGCTGCGGCTTAAAGAAACCAAGTCCAAAGTCACT 2700  
DB 2760 GAATGTGCAACATTTCTGCCCAGCTGCGGCTTAAAGAAACCAAGTCCAAAGTCACT 2819  
QY 2701 TTTGAATGTGAACAAAGAGAAATTAAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 2760  
DB 2820 TTTGAATGTGAACAAAGAGAAATTAAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 2879  
QY 2761 CAGACAGTTAATATCAGTGCAGGCTTCTGTTGTTGTCAGAAAGTAAAGCAGTTGAT 2820  
DB 2880 CAGACAGTTAATATCAGTGCAGGCTTCTGTTGTTGTCAGAAAGTAAAGCAGTTGAT 2939  
QY 2821 AATGCAATGATGATTAAGAGAGGCTCTGAGTTTGTATATATCAGTTCAGAGGC 2880  
DB 2940 AATGCAATGATGATTAAGAGAGGCTCTGAGTTTGTATATATCAGTTCAGAGGC 2999  
QY 2881 AAGCAAACTGACATCTACTCCAATTAACATGAGCTTTTACAAACCCATATCTGTATA 2940  
DB 3000 AAGCAAACTGACATCTACTCCAATTAACATGAGCTTTTACAAACCCATATCTGTATA 3059  
QY 2941 CCACCACTTTTCCCATCAAGTCAATTTGTTTAAACCTAATGTAAGAAAAATCTGCTAGAG 3000  
DB 3060 CCACCACTTTTCCCATCAAGTCAATTTGTTTAAACCTAATGTAAGAAAAATCTGCTAGAG 3119  
QY 3001 GAAAGCTTTGAGAGAACTTCAATGTCAGTGAAGAGAAAGTAAAGTAAAGTAAAGTAAAGT 3060  
DB 3120 GAAAGCTTTGAGAGAACTTCAATGTCAGTGAAGAGAAAGTAAAGTAAAGTAAAGTAAAGT 3179  
QY 3061 AGTACAGTGAACCAATTAAGCCTTAATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 3120

```
|||||
Db 3180 AGTACATGAGCACAATTTAGCCGTAATACATTAGAGAAAATGTTTTAAAGAGCCAGC 3239
QY 3121 TCAGCAATATTAATGAAGTAGTTCCAGTACTAATGAAGGGGCTCCAGATTAAATGAA 3180
Db 3240 TCAGCAATATTAATGAAGTAGTTCCAGTACTAATGAAGGGGCTCCAGATTAAATGAA 3239
QY 3181 ATAGGTTCCAGTGTGAAGAACATTCAGCAGACTAGTAGAAACAGAGGGCCAAAATG 3240
Db 3300 ATAGGTTCCAGTGTGAAGAACATTCAGCAGACTAGTAGAAACAGAGGGCCAAAATG 3359
QY 3241 AATGCTATGCTTAATTAAGGGGTTTTGCAACCTGAGTCTAATAACAAGTCTTCGGA 3300
Db 3360 AATGCTATGCTTAATTAAGGGGTTTTGCAACCTGAGTCTAATAACAAGTCTTCGGA 3419
QY 3301 AGTATATGTAAGCATCTCCGAAATATAAAGCAGAAATATGAAGAGTAGTTCAAGCTGT 3360
Db 3420 AGTATATGTAAGCATCTCCGAAATATAAAGCAGAAATATGAAGAGTAGTTCAAGCTGT 3479
QY 3361 AATACAGATTTCTCCATATCTGATTTAGATTAAGTAAACAGCCATGGAAGTAGT 3420
Db 3480 AATACAGATTTCTCCATATCTGATTTAGATTAAGTAAACAGCCATGGAAGTAGT 3539
QY 3421 CATGCTATCTGAGGTTGTTCTGAGACACCTGATACCTGTTAGATGATGGAATTAAG 3480
Db 3540 CATGCTATCTGAGGTTGTTCTGAGACACCTGATACCTGTTAGATGATGGAATTAAG 3599
QY 3481 GAAGATCTAGTTTGGTGAAGAAATGACATTAAAGAAAGTTGCTGTTTTAGCAAAAGC 3540
Db 3600 GAAGATCTAGTTTGGTGAAGAAATGACATTAAAGAAAGTTGCTGTTTTAGCAAAAGC 3659
QY 3541 GTCCAGAAAGAGAGCTTAGCAGAGAGTCTAGCCCTTTCACCCATACACATTTGGCTAG 3600
Db 3660 GTCCAGAAAGAGAGCTTAGCAGAGAGTCTAGCCCTTTCACCCATACACATTTGGCTAG 3719
QY 3601 GGTTACCGAAGAGGGGCCAAGAAATTAAGAGTCTCTCAAGAGAACTTATCTAGAGAT 3660
Db 3720 GGTTACCGAAGAGGGGCCAAGAAATTAAGAGTCTCTCAAGAGAACTTATCTAGAGAT 3779
QY 3661 GAAGAGCTCCCTCTCCAACTGTTATTTGGTAAAGTAAACAAATATACCTCTCAG 3720
Db 3780 GAAGAGCTCCCTCTCCAACTGTTATTTGGTAAAGTAAACAAATATACCTCTCAG 3839
QY 3721 TCTACTAGGAGTAGACACGTTGCTACGAGTGTCTGCTAAGAACACAGAGAGAAATTA 3780
Db 3840 TCTACTAGGAGTAGACACGTTGCTACGAGTGTCTGCTAAGAACACAGAGAGAAATTA 3899
QY 3781 TTATCATTTGAAGATTAAGTAAATGAGTACAGTAACAGATTAATTTGGCAAGGCAATCT 3840
Db 3900 TTATCATTTGAAGATTAAGTAAATGAGTACAGTAACAGATTAATTTGGCAAGGCAATCT 3959
QY 3841 CAGGACATCACTTACTAGAGAAACAAATGTTCTGCTAGCTGTTTCTCCAGATGC 3900
Db 3960 CAGGACATCACTTACTAGAGAAACAAATGTTCTGCTAGCTGTTTCTCCAGATGC 4019
QY 3901 AGTGAATTTGGAAGACTGAGTGCAGAAATACAAACCCAGAGATCTTCTTGATTTGTTCT 3960
Db 4020 AGTGAATTTGGAAGACTGAGTGCAGAAATACAAACCCAGAGATCTTCTTGATTTGTTCT 4079
QY 3961 TCCAAAATAATGAGGATCAAGTCTGAAAGCCAGGAGATTTGCTGAGTGCACAGAAATG 4020
Db 4080 TCCAAAATAATGAGGATCAAGTCTGAAAGCCAGGAGATTTGCTGAGTGCACAGAAATG 4139
QY 4021 GTTTCAGATGATGAAGAAAGAGAGGGGCTTGGAAGAAATATATCAAGAAAGCAAGC 4080
Db 4140 GTTTCAGATGATGAAGAAAGAGAGGGGCTTGGAAGAAATATATCAAGAAAGCAAGC 4199
QY 4081 ATGAGATTAACACTTAGTGAAGCAGCATCTGGGTGTGAGAGTGAAGAACAGGCTCTGAA 4140
Db 4200 ATGAGATTAACACTTAGTGAAGCAGCATCTGGGTGTGAGAGTGAAGAACAGGCTCTGAA 4259
QY 4141 GAGTGTCTAGGGCTATCTCTCAGAGTGCATTTAAACCACTCAGCAGAGGATACCATG 4200
|||||

Db 4260 GAGTGTCTAGGGCTATCTCTCAGAGTGCATTTAAACCACTCAGCAGAGGATACCATG 4319
QY 4201 CAACATAACTGATAAAGCTCCAGCAGAGAAATGGCTGAAGTGAAGCTGTGTAAGACAG 4260
Db 4320 CAACATAACTGATAAAGCTCCAGCAGAGAAATGGCTGAAGTGAAGCTGTGTAAGACAG 4379
QY 4261 CATGGAGCCAGCCCTTCTAACAGCTACCCCTCATCATTAAGTGAAGTCTTTCGCCCTGAG 4320
Db 4380 CATGGAGCCAGCCCTTCTAACAGCTACCCCTCATCATTAAGTGAAGTCTTTCGCCCTGAG 4439
QY 4321 GAGCTGGGAATCCAGAACAAAGCAGATCAGAAAAAGCAGTATTAACTTCACAGAAAGT 4380
Db 4440 GAGCTGGGAATCCAGAACAAAGCAGATCAGAAAAAGCAGTATTAACTTCACAGAAAGT 4499
QY 4381 AGTGAATACCTTATAGCCAGAAATCCAGAAAGCCCTTCTGCTGCAAGTTGAGGTGCT 4440
Db 4500 AGTGAATACCTTATAGCCAGAAATCCAGAAAGCCCTTCTGCTGCAAGTTGAGGTGCT 4559
QY 4441 GCAGATAGTTCTACAGTAAATAAAGAACAGAGTGAAGAAAGTCAATCCCTTCTTAA 4500
Db 4560 GCAGATAGTTCTACAGTAAATAAAGAACAGAGTGAAGAAAGTCAATCCCTTCTTAA 4619
QY 4501 TGCCCATCATTAAGTATGATAGTGTGATGACATGCAGATTGCTCTGGAGTCTTCAGAAATGA 4560
Db 4620 TGCCCATCATTAAGTATGATAGTGTGATGACATGCAGATTGCTCTGGAGTCTTCAGAAATGA 4679
QY 4561 AACTAACCATCTCAAGAGAGAGCTCATTAAGGTTGTGATGTCGAGAGGACACACCTGGAA 4620
Db 4680 AACTAACCATCTCAAGAGAGAGCTCATTAAGGTTGTGATGTCGAGAGGACACACCTGGAA 4739
QY 4621 GAGTCTGGGCCACACAGATTTTGACGGAACATCTTACTTGGCCAGAGCAAGATCTAGAGGA 4680
Db 4740 GAGTCTGGGCCACACAGATTTTGACGGAACATCTTACTTGGCCAGAGCAAGATCTAGAGGA 4799
QY 4681 ACCCCTTACTGGAATCTGGAATAGCTCTTCTCTGATGACCTGAATCTGATCTTCT 4740
Db 4800 ACCCCTTACTGGAATCTGGAATAGCTCTTCTCTGATGACCTGAATCTGATCTTCT 4859
QY 4741 GAAGACAGAGCCCGAGAGTGTGCTGTTGGCAACATTAACCTTCAACCTGCAATG 4800
Db 4860 GAAGACAGAGCCCGAGAGTGTGCTGTTGGCAACATTAACCTTCAACCTGCAATG 4919
QY 4801 AAAGTCCCAATTTGAAGTTGCAAGATCTGCCCCAGAGTCCAGCTGCTCATACTACT 4860
Db 4920 AAAGTCCCAATTTGAAGTTGCAAGATCTGCCCCAGAGTCCAGCTGCTCATACTACT 4979
QY 4861 GATACCTGCTGGGTATTAATGCAATGGAAGAAAGTGTGAGCAGAGGAGAAACCAAGTTGACA 4920
Db 4980 GATACCTGCTGGGTATTAATGCAATGGAAGAAAGTGTGAGCAGAGGAGAAACCAAGTTGACA 5039
QY 4921 GCTTCACAGAAAGGGTCAACAAAGAAATGTCATGGTGGTGTGCTGGCCAGCCAGAA 4980
Db 5040 GCTTCACAGAAAGGGTCAACAAAGAAATGTCATGGTGGTGTGCTGGCCAGCCAGAA 5099
QY 4981 GAAATTTATGCTGTGTACAAAGTTTGCCAGAAACCAACATCACTTTAACTAATCTAAT 5040
Db 5100 GAAATTTATGCTGTGTACAAAGTTTGCCAGAAACCAACATCACTTTAACTAATCTAAT 5159
QY 5041 ACTGAAGAGACTACTCATGTTGTTATGAAAAACAGATCTGAGTTGTGTGAAACGAGCA 5100
Db 5160 ACTGAAGAGACTACTCATGTTGTTATGAAAAACAGATCTGAGTTGTGTGAAACGAGCA 5219
QY 5101 CTGAAATATTTCTAGGAATTTGGGGAGGAAATGGGTAGTTGCTATTTCTGGGTACC 5160
Db 5220 CTGAAATATTTCTAGGAATTTGGGGAGGAAATGGGTAGTTGCTATTTCTGGGTACC 5279
QY 5161 CAGTCTATTAAGAAAGAAATGCTGAATGAGCATATTTTGAAGTCAAGAGAGATGTG 5220
Db 5280 CAGTCTATTAAGAAAGAAATGCTGAATGAGCATATTTTGAAGTCAAGAGAGATGTG 5339
QY 5221 GTCATGGAAGAAACCAAGAGTCCAAAGCAGACAGAGAAATCCAGAGACAGAAAGATC 5280
Db 5340 GTCATGGAAGAAACCAAGAGTCCAAAGCAGACAGAGAAATCCAGAGACAGAAAGATC 5399
```

```
QY 5281 TTCAGGGGCTAGAAATCTGTGCTATGGGCCCTTCACCAACATGCCACAGATCAACTG 5340
    |||||||
Db 5400 TTCAGGGGCTAGAAATCTGTGCTATGGGCCCTTCACCAACATGCCACAGATCAACTG 5459
    |||||||
QY 5341 GAATGATGATACAGCTGTGTGCTTCTGTGTGAGAGAGCTTTCATCATTCACCCCTT 5400
    |||||||
Db 5460 GAATGATGATACAGCTGTGTGCTTCTGTGTGAGAGAGCTTTCATCATTCACCCCTT 5519
    |||||||
QY 5401 GGCACAGGTGTCACCCCAATTGTTGTTGACAGCAGATGCTGAGACAGAGACAATGGC 5460
    |||||||
Db 5520 GGCACAGGTGTCACCCCAATTGTTGTTGACAGCAGATGCTGAGACAGAGACAATGGC 5579
    |||||||
QY 5461 TTCATGCAATTGGGACAGATGTGTGAGGACCTGTGTGACCCGAGAGTGTGTGGAC 5520
    |||||||
Db 5580 TTCATGCAATTGGGACAGATGTGTGAGGACCTGTGTGACCCGAGAGTGTGTGGAC 5639
    |||||||
QY 5521 AGTGTAGCACTCTACCAAGTCCAGAGAGCTGAGACCTTACCTGATACCCAGATCCCCAC 5580
    |||||||
Db 5640 AGTGTAGCACTCTACCAAGTCCAGAGAGCTGAGACCTTACCTGATACCCAGATCCCCAC 5699
    |||||||
QY 5581 AGCCACTAC 5589
    |||||||
Db 5700 AGCCACTAC 5708
```

Search completed: June 27, 2003, 01:02:14  
Job time : 13935.7 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 26, 2003, 16:33:51 ; Search time 719.151 Seconds

(without alignments)  
11536.634 Million cell updates/sec

Title: US-09-734-672-3\_COPY\_120\_5708

Perfect score: 5589

Sequence: 1 ATGATTTATCTGCTCTTCG.....AGATCCCCACAGCCACTAC 5589

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1055720 seqs, 742224136 residues 2111440

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_NA:\*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCRT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5587.4	100.0	5711	9	US-09-734-672-3
2	5587.4	100.0	5711	9	US-09-982-828-5
3	5582.6	99.9	5711	9	US-09-734-672-5
4	5582.6	99.9	5711	9	US-09-982-828-3
5	5582.6	99.9	5711	9	US-10-022-819-1
6	5581	99.9	5711	9	US-09-734-672-1
7	5581	99.9	5711	9	US-09-982-828-1
8	364.6	6.5	499	9	US-09-911-904-127
9	175.4	3.1	424	10	US-09-864-761-4552
10	147	2.6	147	10	US-09-864-761-21299
11	121	2.2	121	9	US-09-818-875-661
12	121	2.2	121	9	US-09-818-875-662
13	121	2.2	121	9	US-09-818-875-665
14	121	2.2	121	9	US-09-818-875-666
15	121	2.2	121	9	US-09-818-875-669
16	121	2.2	121	9	US-09-818-875-670
17	121	2.2	121	9	US-09-818-875-673
18	121	2.2	121	9	US-09-818-875-674
19	121	2.2	121	9	US-09-818-875-677

c	20	121	2.2	121	9	US-09-818-875-678	Sequence 678, App
c	21	121	2.2	121	9	US-09-818-875-681	Sequence 681, App
c	22	121	2.2	121	9	US-09-818-875-682	Sequence 682, App
c	23	121	2.2	121	9	US-09-818-875-685	Sequence 685, App
c	24	121	2.2	121	9	US-09-818-875-686	Sequence 686, App
c	25	121	2.2	121	9	US-09-818-875-689	Sequence 689, App
c	26	121	2.2	121	9	US-09-818-875-690	Sequence 690, App
c	27	121	2.2	121	9	US-09-818-875-693	Sequence 693, App
c	28	121	2.2	121	9	US-09-818-875-694	Sequence 694, App
c	29	121	2.2	121	9	US-09-818-875-697	Sequence 697, App
c	30	121	2.2	121	9	US-09-818-875-698	Sequence 698, App
c	31	121	2.2	121	9	US-09-818-875-701	Sequence 701, App
c	32	121	2.2	121	9	US-09-818-875-702	Sequence 702, App
c	33	121	2.2	121	9	US-09-818-875-705	Sequence 705, App
c	34	121	2.2	121	9	US-09-818-875-706	Sequence 706, App
c	35	121	2.2	121	9	US-09-818-875-709	Sequence 709, App
c	36	121	2.2	121	9	US-09-818-875-710	Sequence 710, App
c	37	121	2.2	121	9	US-09-818-875-713	Sequence 713, App
c	38	121	2.2	121	9	US-09-818-875-714	Sequence 714, App
c	39	121	2.2	121	9	US-09-818-875-717	Sequence 717, App
c	40	121	2.2	121	9	US-09-818-875-718	Sequence 718, App
c	41	121	2.2	121	9	US-09-818-875-721	Sequence 721, App
c	42	121	2.2	121	9	US-09-818-875-722	Sequence 722, App
c	43	121	2.2	121	9	US-09-818-875-725	Sequence 725, App
c	44	121	2.2	121	9	US-09-818-875-726	Sequence 726, App
c	45	121	2.2	121	9	US-09-818-875-729	Sequence 729, App

## ALIGNMENTS

RESULT 1  
US-09-734-672-3  
; Sequence 3, Application US/09734672  
; Publication No. US20020183268A1  
GENERAL INFORMATION:  
APPLICANT: Murphy, Patricia D.  
Alvares, Christophe C.  
Critz, Brenda S.  
Olson, Sheri J.  
Schelter, Denise B.  
Zeng, Bin  
TITLE OF INVENTION: Coding Sequences of the Human  
BCR1 Gene  
NUMBER OF SEQUENCES: 72  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morgan Lewis & Bockius LLP  
STREET: 1111 Pennsylvania Ave., N.W.  
CITY: Washington  
STATE: District of Columbia  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/734,672  
FILING DATE: 03-Dec-2000  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/966,436  
FILING DATE: 07-No. US20020183268A1-97  
APPLICATION NUMBER: US 08/598,591  
FILING DATE: 12-Feb-96  
ATTORNEY/AGENT INFORMATION:  
NAME: Michael S. Tuscan  
REGISTRATION NUMBER: 43,210  
REFERENCE/DOCKET NUMBER: 44921-5055-02-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-739-3000  
TELEFAX: 202-739-3001

```

: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
:   LENGTH: 5711 base pairs
:   TYPE: nucleic acid
:   STRANDEDNESS: No. US20020183268A1 Relevant
:   TOPOLOGY: Linear
:   MOLECULE TYPE: cdna
:   ORIGINAL SOURCE:
:     ORGANISM: Homo sapiens
:     STRAIN: BRCAl
:   POSITION IN GENOME:
:     CHROMOSOME/SEGMENT: 17
:     MAP POSITION: 17q21
:   SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-734-672-3

Query Match      100.0%; Score 5587.4; DB 9; Length 5711;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5588; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ATGATTTATTCCTCTCTGCTTGAAGAAGTACAAATGTCATTAATGCTATGCAGAAA 60
DB 120 ATGATTTATTCCTCTCTGCTTGAAGAAGTACAAATGTCATTAATGCTATGCAGAAA 179
OY 61 ATCTTAGAGTCCCATCTGCTGAGGTTGATCAAGGAACCTGTCTCCACAAAGTGAC 120
DB 180 ATCTTAGAGTCCCATCTGCTGAGGTTGATCAAGGAACCTGTCTCCACAAAGTGAC 239
OY 121 CACATATTTTGCAATTTTGCAATGCTGAACTTCTACCCAGAAAGAAAGGCCCTTCACAG 180
DB 240 CACATATTTTGCAATTTTGCAATGCTGAACTTCTACCCAGAAAGAAAGGCCCTTCACAG 299
OY 181 TGTCTTTATGTAGATGATATTAACCAAAAGAGCCCTACAGAAATAGAGATTAGT 240
DB 300 TGTCTTTATGTAGATGATATTAACCAAAAGAGCCCTACAGAAATAGAGATTAGT 359
OY 241 CAACCTGTTGAAGAGCTATTTGAAATCATTGTGCTTTTCAGCTTGACACAGGTTGGAG 300
DB 360 CAACCTGTTGAAGAGCTATTTGAAATCATTGTGCTTTTCAGCTTGACACAGGTTGGAG 419
OY 301 TATGCAAAAGCATATTTTGGCAAAAAGAAATTAACCTCTCTGAACTCTTAAAGAT 360
DB 420 TATGCAAAAGCATATTTTGGCAAAAAGAAATTAACCTCTCTGAACTCTTAAAGAT 479
OY 361 GAAGTTTCTATCATCCAAAGTATGAGGCTACAGAAACCGTGCACAAAGACTTCTACAGAT 420
DB 480 GAAGTTTCTATCATCCAAAGTATGAGGCTACAGAAACCGTGCACAAAGACTTCTACAGAT 539
OY 421 GAACCCGAAAATCTCTCTGAGAAACCAAGTCTCAGTGTCAACTCTTAACTTTGA 480
DB 540 GAACCCGAAAATCTCTCTGAGAAACCAAGTCTCAGTGTCAACTCTTAACTTTGA 539
OY 481 ACTGTGAGAACTCTGAGCAAAAGAGCGGATCAACCTCAAAAGAGCTGTGCTACAT 540
DB 600 ACTGTGAGAACTCTGAGCAAAAGAGCGGATCAACCTCAAAAGAGCTGTGCTACAT 659
OY 541 GAATTTGGATCTGATCTCTGAGAGTACCGTTAATTAAGCAACTTATTGACAGTGGGA 600
DB 660 GAATTTGGATCTGATCTCTGAGAGTACCGTTAATTAAGCAACTTATTGACAGTGGGA 719
OY 601 GATCAAGAATGTTTACAAATCACCCCTCAAGGAACCAAGGATGAAATCAAGTTGGATTCT 660
DB 720 GATCAAGAATGTTTACAAATCACCCCTCAAGGAACCAAGGATGAAATCAAGTTGGATTCT 779
OY 661 GAAAAAAGGCTGCTGTGTAATTTTCTGAGACGAGATTAACAATTAATGACATCA 720
DB 780 GAAAAAAGGCTGCTGTGTAATTTTCTGAGACGAGATTAACAATTAATGACATCA 839
OY 721 CCCAGTAATTAATGATTTGGAACACCACTGAGAAAGCTGAGAGGCACTCCAGAAAAG 780
DB 840 CCCAGTAATTAATGATTTGGAACACCACTGAGAAAGCTGAGAGGCACTCCAGAAAAG 899
OY 781 TATCAGGGTAGTTCGTCTTCAAACTTGATGTGAGCCATGTGGCAAAATTAATCTCATGCC 840
```

```

DB 900 TATCAGGGTAGTTCGTCTTCAAACTTGATGTGAGCCATGTGGCAAAATTAATCTCATGCC 959
OY 841 AGCTCATTAAGCATGAGAACACACACACTTTATTACTACTAAGACAGATGATATAGAA 900
DB 960 AGCTCATTAAGCATGAGAACACACACACTTTATTACTACTAAGACAGATGATATAGAA 1019
OY 901 AAGCTGAATTCGTATATAAAGCAAAACAGCCCTGGCTGTGCAAGAGCCCAATTAACGA 960
DB 1020 AAGCTGAATTCGTATATAAAGCAAAACAGCCCTGGCTGTGCAAGAGCCCAATTAACGA 1079
OY 961 TGGGCTGGAAGTAAAGAAACATGTAATGATAGCGGACTCCAGACAGAAAAAAGGTA 1020
DB 1080 TGGGCTGGAAGTAAAGAAACATGTAATGATAGCGGACTCCAGACAGAAAAAAGGTA 1139
OY 1021 GATCTGAATTCGATATCCCTGTGTGAGAGAAAAAAGATGGAATTAAGCAAACTGTCATGC 1080
DB 1140 GATCTGAATTCGATATCCCTGTGTGAGAGAAAAAAGATGGAATTAAGCAAACTGTCATGC 1199
OY 1081 TCAGAGAACTCTAGATATCTGAGAGATGCTTCTTGATTAACATAATAGCAGATTGAG 1140
DB 1200 TCAGAGAACTCTAGATATCTGAGAGATGCTTCTTGATTAACATAATAGCAGATTGAG 1259
OY 1141 AAGGTTAAGAGTGGTTTTCAGAGATGATGAATGTTAGTTCTGATGACTACATGAT 1200
DB 1260 AAGGTTAAGAGTGGTTTTCAGAGATGATGAATGTTAGTTCTGATGACTACATGAT 1319
OY 1201 GGGGAGCTGCAATCAAAATGCAAAAGTAGCTGATGTGAGCTGTAATAGGTAGTACAT 1260
DB 1320 GGGGAGCTGCAATCAAAATGCAAAAGTAGCTGATGTGAGCTGTAATAGGTAGTACAT 1379
OY 1261 GAATATTCCTGTTCTTCAGAGAAATAGACTTACTGGCCAGTATCTCATGAGCCTTTA 1320
DB 1380 GAATATTCCTGTTCTTCAGAGAAATAGACTTACTGGCCAGTATCTCATGAGCCTTTA 1439
OY 1321 ATATGTAAAGGTAAAGAGATTCTCAATCAATCAGTAGAAGTAATTAAGAAACAAATA 1380
DB 1440 ATATGTAAAGGTAAAGAGATTCTCAATCAATCAGTAGAAGTAATTAAGAAACAAATA 1499
OY 1381 TTTGGGAAAACCTATGGAAGAAAGGAGGAGCCCTCCCACTTAAGCTTAAGCTTAAGT 1440
DB 1500 TTTGGGAAAACCTATGGAAGAAAGGAGGAGCCCTCCCACTTAAGCTTAAGCTTAAGT 1559
OY 1441 CTAATTTATAGGACATTTGTTTACTAGAGCACAGATTAATCAAGAGAGTCCCTCAAAAT 1500
DB 1560 CTAATTTATAGGACATTTGTTTACTAGAGCACAGATTAATCAAGAGAGTCCCTCAAAAT 1619
OY 1501 AAATTTAAGCGTAAAGGACACCTACATCAGGCTTCTCATCTGAGAGATTTTATCAAGAA 1560
DB 1620 AAATTTAAGCGTAAAGGAGACCTACATCAGGCTTCTCATCTGAGAGATTTTATCAAGAA 1679
OY 1561 GCAGATTTGGCACTTCAAAAAGACCTCTGAAATGATTAATCAGGGAACCTTAACCAAGGAG 1620
DB 1680 GCAGATTTGGCACTTCAAAAAGACCTCTGAAATGATTAATCAGGGAACCTTAACCAAGGAG 1739
OY 1621 CAGAATGCTCAAGTATGATTAATTAATTAAGTGTGATGAGAATTAACCAAAAGGATAT 1680
DB 1740 CAGAATGCTCAAGTATGATTAATTAATTAAGTGTGATGAGAATTAACCAAAAGGATAT 1799
OY 1681 TCTATTCAAGATGAGAAAAATCCTTAACCAATAGATCACTCGAAAAAAGATCTGCTTTC 1740
DB 1800 TCTATTCAAGATGAGAAAAATCCTTAACCAATAGATCACTCGAAAAAAGATCTGCTTTC 1859
OY 1741 AAAAGGAAAGCTGAACCTAATTAAGCGAGCATATTAACCAATTAAGCACTGAAATTAATTC 1800
DB 1860 AAAAGGAAAGCTGAACCTAATTAAGCGAGCATATTAACCAATTAAGCACTGAAATTAATTC 1919
OY 1801 CACAATTTCAAAAGCACTTAAAGAAATAGGCTGAGAGGAAGTCTTCTACCAAGCATATT 1860
DB 1920 CACAATTTCAAAAGCACTTAAAGAAATAGGCTGAGAGGAAGTCTTCTACCAAGCATATT 1979
OY 1861 CATGCCCTTGAACCTAGTACGTAGAAATCTTAAGCCCACTTAATTTGATCTGAATTTGCAA 1920
```

Db 1980 CATGCGTTGMACTAGTACTAGTAGAATCTAAGCCCACTAATTGTACTGATTTGCCAA 2039  
QY 1921 ATTGATAGTTGTTCTAGCAGTGAAGATTAAGAAAAAATGACCAACCAATGCCAGTC 1980  
Db 2040 ATTGATAGTTGTTCTAGCAGTGAAGATTAAGAAAAAATGACCAACCAATGCCAGTC 2099  
QY 1981 AGGCACAGCAGAAACCTACACTCATGGAAGGTAAAGAACCTGCAACCTGAGCCAAAGAG 2040  
Db 2100 AGGCACAGCAGAAACCTACACTCATGGAAGGTAAAGAACCTGCAACCTGAGCCAAAGAG 2159  
QY 2041 AGTAAACAGCCCAATGAACAGACAGATTAAGAACATGACAGTATATCTTCCAGAGCTG 2100  
Db 2160 AGTAAACAGCCCAATGAACAGACAGATTAAGAACATGACAGTATATCTTCCAGAGCTG 2219  
QY 2101 AAGTTAAACAATGACACCTGGTCTTTTACTAGTGTCAATATACAGTACCTTAAAGAA 2160  
Db 2220 AAGTTAAACAATGACACCTGGTCTTTTACTAGTGTCAATATACAGTACCTTAAAGAA 2279  
QY 2161 TTTGTCAATCTAGCCTTCCAGAGAAAGAAAAAGAGAACTAGAAACAGTTAAAGTG 2220  
Db 2280 TTTGTCAATCTAGCCTTCCAGAGAAAGAAAAAGAGAACTAGAAACAGTTAAAGTG 2339  
QY 2221 TCTAATAATGCTGAACACCCCAAGATCTCATGTTAAGTGAAGAGGTTTTCAGAACT 2280  
Db 2340 TCTAATAATGCTGAACACCCCAAGATCTCATGTTAAGTGAAGAGGTTTTCAGAACT 2399  
QY 2281 GAAAGATCTGTAGAGAGTACAGTATTTGATTTGCTAGTGTGATTTAGGCACTCAG 2340  
Db 2400 GAAAGATCTGTAGAGAGTACAGTATTTGATTTGCTAGTGTGATTTAGGCACTCAG 2459  
QY 2341 GAAAGATCTGTAGAGAGTACAGTATTTGATTTGCTAGGAGGCAAAAACAGAACCAATAAA 2400  
Db 2460 GAAAGATCTGTAGAGAGTACAGTATTTGATTTGCTAGGAGGCAAAAACAGAACCAATAAA 2519  
QY 2401 TGTGTAGTACAGTGTGACAGATTTGAAACCCCAAGGAGTATATGATGTTTCCAAA 2460  
Db 2520 TGTGTAGTACAGTGTGACAGATTTGAAACCCCAAGGAGTATATGATGTTTCCAAA 2579  
QY 2461 GATAATAGAAATGACACAGAGGCTTTAAGTATCATTTGGACATGAAGTTAACCCAGT 2520  
Db 2580 GATAATAGAAATGACACAGAGGCTTTAAGTATCATTTGGACATGAAGTTAACCCAGT 2639  
QY 2521 CGGGAACACAGCATAGAAATGGAAGAGTGAACCTTGATGCTCAGTATTTGCGAATATAC 2580  
Db 2640 CGGGAACACAGCATAGAAATGGAAGAGTGAACCTTGATGCTCAGTATTTGCGAATATAC 2699  
QY 2581 TTCAAGTTCGAAGGCGCAGTATTTGCTGTCTTTTCAATCCAGAAATGCGAAGAG 2640  
Db 2700 TTCAAGTTCGAAGGCGCAGTATTTGCTGTCTTTTCAATCCAGAAATGCGAAGAG 2759  
QY 2641 GAATGTGCAACATCTCTGCCACCTGGGCTCTTAAAGAAACCAAGTCCAAAGTCAGT 2700  
Db 2760 GAATGTGCAACATCTCTGCCACCTGGGCTCTTAAAGAAACCAAGTCCAAAGTCAGT 2819  
QY 2701 TTTGAATGTGAACAAAGAGAGAAATCAAGAAAGATGAGTCTAATATCAAGCCTGTA 2760  
Db 2820 TTTGAATGTGAACAAAGAGAGAAATCAAGAAAGATGAGTCTAATATCAAGCCTGTA 2879  
QY 2761 CAGACATGTAATATCTCTGAGGCTTCTGTGTGTGTGTCAGAAAGATAGCCAGTTGAT 2820  
Db 2880 CAGACATGTAATATCTCTGAGGCTTCTGTGTGTGTGTCAGAAAGATAGCCAGTTGAT 2939  
QY 2821 AATGCCAATGTAGTATCAAGAGAGGCTTAGGTTTGTCTATCATCTAGTTCAGAGGC 2880  
Db 2940 AATGCCAATGTAGTATCAAGAGAGGCTTAGGTTTGTCTATCATCTAGTTCAGAGGC 2999  
QY 2881 AAGCAACTGGAATCTATCTCCAAATTAACATGACTTTTACAAAACCATATCTGTATA 2940  
Db 3000 AAGCAACTGGAATCTATCTCCAAATTAACATGACTTTTACAAAACCATATCTGTATA 3059  
QY 2941 CCACCACTTTTCCCATCAAGTCATTTGTTAAACCTAATGTAAAGAAATCTGCTAGAG 3000  
Db 3060 CCACCACTTTTCCCATCAAGTCATTTGTTAAACCTAATGTAAAGAAATCTGCTAGAG 3119

QY 3001 GAAACCTTTGAGGACATTTCAATGTGCACCTGAAGAGAAATGGGAAATGAGAACATTTCCA 3060  
Db 3120 GAAACCTTTGAGGACATTTCAATGTGCACCTGAAGAGAAATGGGAAATGAGAACATTTCCA 3179  
QY 3061 AGTACAGTGAGCACAATTTAGCCGTAATACATTAGAGAAAAATGTTTTAAAGAACCCAGC 3120  
Db 3180 AGTACAGTGAGCACAATTTAGCCGTAATACATTAGAGAAAAATGTTTTAAAGAACCCAGC 3239  
QY 3121 TCAGCAATATTTATGAAGTGTTCAGTACTAATGAAGTGGGCTCCAGTATTAATGAA 3180  
Db 3240 TCAGCAATATTTATGAAGTGTTCAGTACTAATGAAGTGGGCTCCAGTATTAATGAA 3299  
QY 3181 ATAGGTTCCAGTATGAAGAAATTTCAAGCAGAACTAGTGAAGAACAGAGGCCAAATTTG 3240  
Db 3300 ATAGGTTCCAGTATGAAGAAATTTCAAGCAGAACTAGTGAAGAACAGAGGCCAAATTTG 3359  
QY 3241 AATGCTATGCTTAGATTAAGGGGTTTTCACACCTGAGGCTATTAACAAAGCTTCTCGA 3300  
Db 3360 AATGCTATGCTTAGATTAAGGGGTTTTCACACCTGAGGCTATTAACAAAGCTTCTCGA 3419  
QY 3301 AGTAAATTTAGCATCTCTGAATAAAAAGCAGAAATATGAAGATAGTTTCAGACTGTT 3360  
Db 3420 AGTAAATTTAGCATCTCTGAATAAAAAGCAGAAATATGAAGATAGTTTCAGACTGTT 3479  
QY 3361 AATACAGATTTCTCTCATATCTGATTTTCAGATTAAGTACAGCCTATGAGGAGTATG 3420  
Db 3480 AATACAGATTTCTCTCATATCTGATTTTCAGATTAAGTACAGCCTATGAGGAGTATG 3539  
QY 3421 CATGATCTTCAGGTTTGTCTGAGACACCTGATGACCTGTTAGATGATGTAAGTAAG 3480  
Db 3540 CATGATCTTCAGGTTTGTCTGAGACACCTGATGACCTGTTAGATGATGTAAGTAAG 3599  
QY 3481 GAAATACATAGTTTCTCTGAATTAAGCAATTAAGAAAGTCTGCTTTTACCAAAAGC 3540  
Db 3600 GAAATACATAGTTTCTCTGAATTAAGCAATTAAGAAAGTCTGCTTTTACCAAAAGC 3659  
QY 3541 GTCCAGAAAGAGAGCTTAGCAGAGTCCAGGCTTTTACCCATACACATTTGGGCTCAG 3600  
Db 3660 GTCCAGAAAGAGAGCTTAGCAGAGTCCAGGCTTTTACCCATACACATTTGGGCTCAG 3719  
QY 3601 GGTTCACGAAGAGGGGCCAAGAAATTAAGAGTCTCTCAGAAAGAGAACTTATCTAGAGAT 3660  
Db 3720 GGTTCACGAAGAGGGGCCAAGAAATTAAGAGTCTCTCAGAAAGAGAACTTATCTAGAGAT 3779  
QY 3661 GAAGAGCTTCCCTGCTTCCAAACCTGTTATTTGGTAAAGTAAACAATATACCTTCTCAG 3720  
Db 3780 GAAGAGCTTCCCTGCTTCCAAACCTGTTATTTGGTAAAGTAAACAATATACCTTCTCAG 3839  
QY 3721 TCTATAGGATGACACCGTGTGTACAGAGTGTCTGTAAAGAACAGAGGAGAAATTTA 3780  
Db 3840 TCTATAGGATGACACCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3899  
QY 3781 TTATCATTTGAAGATAGCTTAATGAATGACTGACAGTAAACAGGTAATTTGGCAAGGATCT 3840  
Db 3900 TTATCATTTGAAGATAGCTTAATGAATGACTGACAGTAAACAGGTAATTTGGCAAGGATCT 3959  
QY 3841 CAGGAACATCACCCTTAGTGAAGAAACAAATGTTCTGCTAGCTGTTTCTTACAGTGC 3900  
Db 3960 CAGGAACATCACCCTTAGTGAAGAAACAAATGTTCTGCTAGCTGTTTCTTACAGTGC 4019  
QY 3901 AGTGAATTTGAAGATCTGACTGCAAAATACAAACACCCAGAGATCTTCTGATTTGGTCT 3960  
Db 4020 AGTGAATTTGAAGATCTGACTGCAAAATACAAACACCCAGAGATCTTCTGATTTGGTCT 4079  
QY 3961 TCCAAACCAATGAGGATCAGTCTGAAGCCAGGAGTTGCTGTGATGACAAAGAAATG 4020  
Db 4080 TCCAAACCAATGAGGATCAGTCTGAAGCCAGGAGTTGCTGTGATGACAAAGAAATG 4139  
QY 4021 GTTTTCAGATGATGAAGAAAGAGAGAGGCTTGGAGAAATTAATGAAGAGGCAAGC 4080  
Db 4140 GTTTTCAGATGATGAAGAAAGAGAGGCTTGGAGAAATTAATGAAGAGGCAAGC 4199

QY 4081 ATGATTCAAACCTTAGTGAGACAGCATCTGGGTGAGAGTGAAACAAGCTCTCTGAA 4140  
 |||||||  
 Db 4200 ATGGATTCAAACCTTAGTGAGACAGCATCTGGGTGAGAGTGAAACAAGCTCTCTGAA 4259  
 |||||||  
 QY 4141 GAGTCTCAGAGGCTTCCCTCAGAGTGACATTTTACCCTCAGACAGAGGATTCAGG 4200  
 |||||||  
 Db 4260 GACTGCTCAGAGGCTTCCCTCAGAGTGACATTTTACCCTCAGACAGGATTCAGG 4319  
 |||||||  
 QY 4201 CAACATAACCTGATTAACCTCCAGCAGGAAATGGCTGAATAGAAAGCTGTGTAACAG 4260  
 |||||||  
 Db 4320 CAACATAACCTGATTAACCTCCAGCAGGAAATGGCTGAATAGAAAGCTGTGTAACAG 4379  
 |||||||  
 QY 4261 CATGGAGACCAAGCTTCTTAACAGCTACCTTCCATCATTAAGTACTTCTGCTTGAG 4320  
 |||||||  
 Db 4380 CATGGAGACCAAGCTTCTTAACAGCTACCTTCCATCATTAAGTACTTCTGCTTGAG 4439  
 |||||||  
 QY 4321 GACCTGCGCAAAATCCAGAACAAAGCAGATCAGAAAAAGCAGTATTAACTCAGAGAAAGT 4380  
 |||||||  
 Db 4440 GACCTGCGCAAAATCCAGAACAAAGCAGATCAGAAAAAGCAGTATTAACTCAGAGAAAGT 4499  
 |||||||  
 QY 4381 AGTGAATACCTTATTAAGCCAGAAATCCAGAAAGGCTTCTGCTGACAAAGTTGAGGTGCT 4440  
 |||||||  
 Db 4500 AGTGAATACCTTATTAAGCCAGAAATCCAGAAAGGCTTCTGCTGACAAAGTTGAGGTGCT 4559  
 |||||||  
 QY 4441 GCAGATATCTTACAGATTAATAAATAAGAACAGAGTGGAAGGTGATCCCTCTTAA 4500  
 |||||||  
 Db 4560 GCAGATATCTTACAGATTAATAAATAAGAACAGAGTGGAAGGTGATCCCTCTTAA 4619  
 |||||||  
 QY 4501 TGCCCATCATTAAGATGATAGGTGATGATGACAGTGTGCTGGAGTTCAGATATA 4560  
 |||||||  
 Db 4620 TGCCCATCATTAAGATGATAGGTGATGATGACAGTGTGCTGGAGTTCAGATATA 4679  
 |||||||  
 QY 4561 AACTACCCATCTCAAGAGAGAGCTCATTAAGTTGTTGATGAGAGAGACAGCTGAA 4620  
 |||||||  
 Db 4680 AACTACCCATCTCAAGAGAGAGCTCATTAAGTTGTTGATGAGAGAGACAGCTGAA 4739  
 |||||||  
 QY 4621 GAGTCTGGGCCACAGATTTTGACGGAACATCTTACTTGCCAAAGCAGATCTAGAGGA 4680  
 |||||||  
 Db 4740 GAGTCTGGGCCACAGATTTTGACGGAACATCTTACTTGCCAAAGCAGATCTAGAGGA 4799  
 |||||||  
 QY 4681 ACCCTTACCTGGAATCTGAATCAGCTCTTCTGATGACCTGAATCTGATCTCT 4740  
 |||||||  
 Db 4800 ACCCTTACCTGGAATCTGAATCAGCTCTTCTGATGACCTGAATCTGATCTCT 4859  
 |||||||  
 QY 4741 GAAGACAGAGCCCAAGAGTCAAGCTGCTGTGGCAACATACATCTTCAACTCTGCATG 4800  
 |||||||  
 Db 4860 GAAGACAGAGCCCAAGAGTCAAGCTGCTGTGGCAACATACATCTTCAACTCTGCATG 4919  
 |||||||  
 QY 4801 AAAGTCCCAATGAAGTTGCAAGATCTGCCCCAGAGTCCAGCTGCTGCTACTACT 4860  
 |||||||  
 Db 4920 AAAGTCCCAATGAAGTTGCAAGATCTGCCCCAGAGTCCAGCTGCTGCTACTACT 4979  
 |||||||  
 QY 4861 GATACCTGCTGGTATTAATGCAATGGAAGAAAGTGAGAGAGGAGAGCAAAATGACA 4920  
 |||||||  
 Db 4980 GATACCTGCTGGTATTAATGCAATGGAAGAAAGTGAGAGAGGAGAGCAAAATGACA 5039  
 |||||||  
 QY 4921 GCTTCACAGAAAGGCTCAACAAAAGAAATGTCATGCTGCTGCTGAGCCCGACAGA 4980  
 |||||||  
 Db 5040 GCTTCACAGAAAGGCTCAACAAAAGAAATGTCATGCTGCTGCTGAGCCCGACAGA 5099  
 |||||||  
 QY 4981 GAATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5040  
 |||||||  
 Db 5100 GAATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5159  
 |||||||  
 QY 5041 ACTGAAGAGACTACTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5100  
 |||||||  
 Db 5160 ACTGAAGAGACTACTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5219  
 |||||||  
 QY 5101 CTGAATATATTTTCTAGAAATGCGGAGAGAAATGGGTAGTTAGCTATTTCTGGGTGAC 5160  
 |||||||  
 Db 5220 CTGAATATATTTTCTAGAAATGCGGAGAGAAATGGGTAGTTAGCTATTTCTGGGTGAC 5279  
 |||||||  
 QY 5161 CAGTCTATTTAAAGAAAGAAATGCTGAATGACATGATTTTGAAGTCAAGAGAGATG 5220  
 |||||||

Db 5280 CAGTCTATTTAAAGAAAGAAATGCTGAATGAGCATGATTTTGAAGTCAAGGAGATG 5339  
 |||||||  
 QY 5221 GTCATGCAAGAAACACCAAGGTCCAAAGCAGAGCAAGATCCAGAGCAAGATC 5280  
 |||||||  
 Db 5340 GTCATGCAAGAAACACCAAGGTCCAAAGCAGAGCAAGATCCAGAGCAAGATC 5399  
 |||||||  
 QY 5281 TTCAGGGGCTAGAAATCTGTTGCTATGAGGCCCTTCCACCAACATGCCAGATCAATG 5340  
 |||||||  
 Db 5400 TTCAGGGGCTAGAAATCTGTTGCTATGAGGCCCTTCCACCAACATGCCAGATCAATG 5459  
 |||||||  
 QY 5341 GAATGATGCTACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5400  
 |||||||  
 Db 5460 GAATGATGCTACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5519  
 |||||||  
 QY 5401 GGCACAGGTGTCACCCCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5460  
 |||||||  
 Db 5520 GGCACAGGTGTCACCCCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5579  
 |||||||  
 QY 5461 TTCATGCAATTTGGCAGATGCTGAGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5520  
 |||||||  
 Db 5580 TTCATGCAATTTGGCAGATGCTGAGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5639  
 |||||||  
 QY 5521 AGTGTAGCACTTACCATGCTCCAGAGAGCTGACACACTACCTGATACCCAGATCCCT 5580  
 |||||||  
 Db 5640 AGTGTAGCACTTACCATGCTCCAGAGAGCTGACACACTACCTGATACCCAGATCCCT 5699  
 |||||||  
 QY 5581 AGCCACTAC 5589  
 |||||||  
 Db 5700 AGCCACTAC 5708  
 |||||||

RESULT 2  
 US-09-982-828-5  
 : Sequence 5, Application US/09982828  
 : Publication No. US20030022184A1  
 :  
 : GENERAL INFORMATION:  
 :  
 : APPLICANT: Murphy, Patricia D.  
 : Allen, Antonette C.  
 : Alvarres, Christopher P.  
 : Critz, Brenda S.  
 : Olson, Sheri J.  
 : Thurber, Denise  
 : Zeng, Bin  
 :  
 : TITLE OF INVENTION: Coding Sequences of the Human  
 : BRCA1 Gene  
 :  
 : NUMBER OF SEQUENCES: 72  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Morgan Lewis & Bockius LLP  
 : STREET: 1111 Pennsylvania Avenue N. W.  
 : CITY: Washington  
 : STATE: DC  
 : COUNTRY: USA  
 : ZIP: 20004  
 :  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: Patentin Release #1.0, Version #1.30  
 :  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/09/982,828  
 : FILING DATE: 22-Oct-2001  
 : CLASSIFICATION: <unknown>  
 :  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: US 09/074,453  
 : FILING DATE: 1998-05-06  
 : APPLICATION NUMBER: US 08/798,691  
 : FILING DATE: 1997-02-12  
 : APPLICATION NUMBER: US 08/598,591  
 : FILING DATE: 1996-02-12  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: Michael S. Tuscan  
 : REGISTRATION NUMBER: 43,210



REFERENCE/DOCKET NUMBER: 44921-5053-01-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-739-3000  
TELEFAX: 202-739-3001  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5711 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
STRAIN: BRCA1 (om13)  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: 17  
MAP POSITION: 17q21  
SEQUENCE DESCRIPTION: SEQ ID NO: 5:  
US-09-982-828-5

Query Match 100.0%; Score 5587.4; DB 9; Length 5711;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 5588; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGATTATCTGCTCTTGGCGTTGAAGAATGTCATTAATGCTATGAGAAA 60  
DB ATGGATTATCTGCTCTTGGCGTTGAAGAATGTCATTAATGCTATGAGAAA 179  
QY 61 ATCTTGAGTGTCCCTGCTGCTGAGTGTATCAAGAACCTGTCCCAAGGTGAC 120  
DB 180 ATCTTGAGTGTCCCTGCTGCTGAGTGTATCAAGAACCTGTCCCAAGGTGAC 239  
QY 121 CACATATTTTGGCAATTTTGCATGTAAGACTCTCAACAGAAAGAGGCTTCACAG 180  
DB 240 CACATATTTTGGCAATTTTGCATGTAAGACTCTCAACAGAAAGAGGCTTCACAG 299  
QY 181 TGTCTTTATGTAAGATGATATACCAAAAGAGCTCAAGAAAGTACGAGATTAGT 240  
DB 300 TGTCTTTATGTAAGATGATATACCAAAAGAGCTCAAGAAAGTACGAGATTAGT 359  
QY 241 CAACCTGTGAGAGCTATTTGAAATCATTTTGTCTTTTACGTTTACACAGCTTTGAG 300  
DB 360 CAACCTGTGAGAGCTATTTGAAATCATTTTGTCTTTTACGTTTACACAGCTTTGAG 419  
QY 301 TATGCAACAGCATATTTTGCAGAAAAGAAATACCTCTGCAACATCTAAAGAT 360  
DB 420 TATGCAACAGCATATTTTGCAGAAAAGAAATACCTCTGCAACATCTAAAGAT 479  
QY 361 GAAGTTTCTATCATCAAAAGTATGGCTACAGAAACCGTGCCAAAAGACTTTACAGAGT 420  
DB 480 GAAGTTTCTATCATCAAAAGTATGGCTACAGAAACCGTGCCAAAAGACTTTACAGAGT 539  
QY 421 GAACCCGAAATCCTCTCTGCGAGAAACCGTCTAGTGTCCACTCTTAACCTTGA 480  
DB 540 GAACCCGAAATCCTCTCTGCGAGAAACCGTCTAGTGTCCACTCTTAACCTTGA 599  
QY 481 ACTGTGAGAACTCTGAGACAAAGAGCGGATACAACTCAAAAAGAGCTGTCTACATT 540  
DB 600 ACTGTGAGAACTCTGAGACAAAGAGCGGATACAACTCAAAAAGAGCTGTCTACATT 659  
QY 541 GAATTGGGATCTGANTCTTCTGAAGATACCGTTAATAGGCAACTTATTGCACTGTGGA 600  
DB 660 GAATTGGGATCTGANTCTTCTGAAGATACCGTTAATAGGCAACTTATTGCACTGTGGA 719  
QY 601 GATCAAGATTTGTACAAATCAACCCCTCAAGAACCCAGGATGAATTCAGTTTGATTTCT 660  
DB 720 GATCAAGATTTGTACAAATCAACCCCTCAAGAACCCAGGATGAATTCAGTTTGATTTCT 779  
QY 661 GCAAAAAAGCTGCTGTGAATTTTCTGAGAGGATGTAACAATCTGACATCTCAA 720  
DB 780 GCAAAAAAGCTGCTGTGAATTTTCTGAGAGGATGTAACAATCTGACATCTCAA 839  
QY 721 CCAGTAATATGATTGGAACACACTGAGAACCGTGACGTGAGAGCATCCAGAAAAG 780

DB 840 CCAGTAATATGATTGGAACACACTGAGAACCGTGACACTGAGAGCATCCAGAAAAG 899  
QY 781 TATCAGAGGTAGTCTGTTTCAAACTGCATGTGGAGCATGTGGACAAATATCTCATGCC 840  
DB 900 TATCAGAGGTAGTCTGTTTCAAACTGCATGTGGAGCATGTGGACAAATATCTCATGCC 959  
QY 841 AGCTCATTCAGCATGAGAACACAGCTTTATTAATCTACTCACTAAAGAGAAATGATGAA 900  
DB 960 AGCTCATTCAGCATGAGAACACAGCTTTATTAATCTACTCACTAAAGAGAAATGATGAA 1019  
QY 901 AAGCTGAATTTCTGTAATTAAGCAAAACAGCTGGCTTTAGCAAGAGCCAACTATACAGA 960  
DB 1020 AAGCTGAATTTCTGTAATTAAGCAAAACAGCTGGCTTTAGCAAGAGCCAACTATACAGA 1079  
QY 961 TGGGCTGGAAAGTAAAGAAACATGATAGAGCGGACCTCCGACAGAAAAAGGTA 1020  
DB 1080 TGGGCTGGAAAGTAAAGAAACATGATAGAGCGGACCTCCGACAGAAAAAGGTA 1139  
QY 1021 GATCTGAATGCTGATCCCTGTGTGAGAGAAAGAAATGAATAGCAAGAACTGCATGC 1080  
DB 1140 GATCTGAATGCTGATCCCTGTGTGAGAGAAAGAAATGAATAGCAAGAACTGCATGC 1199  
QY 1081 TCAGAGAAATCTAGAGATGATGAGATGTTCTTGATTAACACTAATATGACCACTTACAG 1140  
DB 1200 TCAGAGAAATCTAGAGATGATGAGATGTTCTTGATTAACACTAATATGACCACTTACAG 1259  
QY 1141 AAGTTAATGAGTGTTTCCAGAACTGATGAACTGTAGCTGTGATGCTACATGAT 1200  
DB 1260 AAGTTAATGAGTGTTTCCAGAACTGATGAACTGTAGCTGTGATGCTACATGAT 1319  
QY 1201 GGGAGCTCTGAATCAATGCAAGTACGATGATGATGATGATGATGATGATGATGATGAT 1260  
DB 1320 GGGAGCTCTGAATCAATGCAAGTACGATGATGATGATGATGATGATGATGATGATGAT 1379  
QY 1261 GAATATCTGCTTCTTACAGAAATAGACTTACTGGCAGTATCCTCATGAGGCTTTA 1320  
DB 1380 GAATATCTGCTTCTTACAGAAATAGACTTACTGGCAGTATCCTCATGAGGCTTTA 1439  
QY 1321 ATATGTAAAGTAAAGAGTTCCTCAATATGATGATGATGATGATGATGATGATGATGAT 1380  
DB 1440 ATATGTAAAGTAAAGAGTTCCTCAATATGATGATGATGATGATGATGATGATGATGAT 1499  
QY 1381 TTTGGGAAAACTATGGAAGAGGCAAGGCTCCCAACTTAAGCCATGATGATGATGAT 1440  
DB 1500 TTTGGGAAAACTATGGAAGAGGCAAGGCTCCCAACTTAAGCCATGATGATGATGATGAT 1559  
QY 1441 CTAAATATAGGACATTTGTTACTGAGCCACAGATATACAGAGGCTCCCTCACAAAT 1500  
DB 1560 CTAAATATAGGACATTTGTTACTGAGCCACAGATATACAGAGGCTCCCTCACAAAT 1619  
QY 1501 AAATTAAGCGTAAAGAGAGCTACATGAGGCTTCATCTGAGATTTTATCAAGAAA 1560  
DB 1620 AAATTAAGCGTAAAGAGAGCTACATGAGGCTTCATCTGAGATTTTATCAAGAAA 1679  
QY 1561 GCAGATTTGGCAGTTCAAAAGACTCTGAATATGATGATGATGATGATGATGATGATGAT 1620  
DB 1680 GCAGATTTGGCAGTTCAAAAGACTCTGAATATGATGATGATGATGATGATGATGATGAT 1739  
QY 1621 CAGAAATGCTCAAGTATGATATTAATAGTGTGATGAGATTAAGCAAAAGGTGAT 1680  
DB 1740 CAGAAATGCTCAAGTATGATATTAATAGTGTGATGAGATTAAGCAAAAGGTGAT 1799  
QY 1681 TCTATTCAAGATGAGAAAAATCTTAACCCCAATAGATCACTGAAAAAGATGCTTTTC 1740  
DB 1800 TCTATTCAAGATGAGAAAAATCTTAACCCCAATAGATCACTGAAAAAGATGCTTTTC 1859  
QY 1741 AAAAGCAAGCTGAACCTTAAGACAGCAGATTAAGCAATATGGAATCTGATTAATATTC 1800  
DB 1860 AAAAGCAAGCTGAACCTTAAGACAGCAGATTAAGCAATATGGAATCTGATTAATATTC 1919  
QY 1801 CACAAATTCAAAAGCACCTTAAGAAAGATAGCTGAGAGAGAGCTTCTACAGAGCATATT 1860

Dd 1920 CACAATTCAAAAACCTAAAAAATAGCTGAGGAGGAGTCTTTCTACAGGACATATTT 1979  
Qy 1861 CATGGGCTTGAACTAGTATGCTGAATATTAAGCCACCTATTTGTAATGCA 1920  
Dd 1980 CATGGGCTTGAACTAGTATGCTGAATATTAAGCCACCTATTTGTAATGCA 2039  
Qy 1921 ATTTAGTATGCTTCTAGCAGTGAAGATTAAGAAAAAAGTACAACTATGCGCAGTC 1980  
Dd 2040 ATTTAGTATGCTTCTAGCAGTGAAGATTAAGAAAAAAGTACAACTATGCGCAGTC 2099  
Qy 1981 AGGCACAGCAGAAACCTTACACTCATGGAAGTAAAGAACTGCAACTGAGCCAGAAAG 2040  
Dd 2100 AGGCACAGCAGAAACCTTACACTCATGGAAGTAAAGAACTGCAACTGAGCCAGAAAG 2159  
Qy 2041 AGTACCAAGCCAAATGAACAGACAGTAAGACATGACATGATCTTTCCAGAGCTG 2100  
Dd 2160 AGTAAACAAAGCCAAATGAACAGACAGTAAGACATGACATGATCTTTCCAGAGCTG 2219  
Qy 2101 AAGTTAACAAATGACACCTGGTTCTTTTACTAAGTGTTCAAATPCCAGTGAATTAAGAA 2160  
Dd 2220 AAGTTAACAAATGACACCTGGTTCTTTTACTAAGTGTTCAAATPCCAGTGAATTAAGAA 2279  
Qy 2161 TTTTGCAATCCTAGCCTTCCAAAGAGAAAAAGAGAACTAGAAACAGTTAAAGTG 2220  
Dd 2280 TTTTGCAATCCTAGCCTTCCAAAGAGAAAAAGAGAACTAGAAACAGTTAAAGTG 2339  
Qy 2221 TCTAATTAATGCTGAAGAGCCCAAGATCTCATGTTAAGTGGAGAAAGGGTTTGGCAACT 2280  
Dd 2340 TCTAATTAATGCTGAAGAGCCCAAGATCTCATGTTAAGTGGAGAAAGGGTTTGGCAACT 2399  
Qy 2281 GAAAGATCTGTAGAGATGACAGATATTTTCATTGGTACCTGATGATTAAGCAGCTCAG 2340  
Dd 2400 GAAAGATCTGTAGAGATGACAGATATTTTCATTGGTACCTGATGATTAAGCAGCTCAG 2459  
Qy 2341 GAAATATCTCGTCTACTGGAATTTAGCACTAGGAGAACCAAAACAAACCAATTA 2400  
Dd 2460 GAAATATCTCGTCTACTGGAATTTAGCACTAGGAGAACCAAAACAAACCAATTA 2519  
Qy 2401 TGTGTGAGTCAAGTGTGACAGATTTTGAACCCCAAGGAGACTAATTCATGGTTGTTCCAA 2460  
Dd 2520 TGTGTGAGTCAAGTGTGACAGATTTTGAACCCCAAGGAGACTAATTCATGGTTGTTCCAA 2579  
Qy 2461 GATTAATGAAATGACACAGAGGCTTTAAGTATCCATTGGGACATGAGTTAAACACAGT 2520  
Dd 2580 GATTAATGAAATGACACAGAGGCTTTAAGTATCCATTGGGACATGAGTTAAACACAGT 2639  
Qy 2521 CGGGAACAGAGCTGAAGATGGAAGAAATGAACTGATGATTTGAGATTA 2580  
Dd 2640 CGGGAACAGAGCTGAAGATGGAAGAAATGAACTGATGATTTGAGATTA 2699  
Qy 2581 TTCAAGGTTTCAAGCGCCAGTCAATTTGCTGTTTCAAAATCCAGGAATCCAGAAAG 2640  
Dd 2700 TTCAAGGTTTCAAGCGCCAGTCAATTTGCTGTTTCAAAATCCAGGAATCCAGAAAG 2759  
Qy 2641 GAATGTGCAACATTTCTGCCCCACCTGCGCTTAAAGAAACAAAGTCCAAAGTCACT 2700  
Dd 2760 GAATGTGCAACATTTCTGCCCCACCTGCGCTTAAAGAAACAAAGTCCAAAGTCACT 2819  
Qy 2701 TTTGAATGTGAACAAAGAAAGAAATCAAGAAAGAAATGAGTCTAATTCAAAGCCTGTA 2760  
Dd 2820 TTTGAATGTGAACAAAGAAAGAAATCAAGAAAGAAATGAGTCTAATTCAAAGCCTGTA 2879  
Qy 2761 CAGACAGTTAATATCTAGCTGAGGCTTTCCTGTTGTTGTCAGAAAGATTAAGCAGTTGAT 2820  
Dd 2880 CAGACAGTTAATATCTAGCTGAGGCTTTCCTGTTGTTGTCAGAAAGATTAAGCAGTTGAT 2939  
Qy 2821 AATGCAAAATGATATCAAAAGAGGCTCTAGGTTTGTCTATCATCTCAGTTCAAGGCG 2880  
Dd 2940 AATGCAAAATGATATCAAAAGAGGCTCTAGGTTTGTCTATCATCTCAGTTCAAGGCG 2999  
Qy 2881 AAGCAAACTGAGCTATTAATCTCAAAATTAACATGAGCTTTTACAAAACCCATATGCTATA 2940  
Dd 3000 AAGCAAACTGAGCTATTAATCTCAAAATTAACATGAGCTTTTACAAAACCCATATGCTATA 3059

Qy 2941 CCACACCTTTTCCCATCAAGTCAATTTGTTAAACCTAAATGTAGAAAAATCTGCTAGAG 3000  
Dd 3060 CCACACCTTTTCCCATCAAGTCAATTTGTTAAACCTAAATGTAGAAAAATCTGCTAGAG 3119  
Qy 3001 GAAACCTTTGAGAACATTTCAATGTGCACCTGAAGAAAGAAATGGAATGAGAACATTTCA 3060  
Dd 3120 GAAACCTTTGAGAACATTTCAATGTGCACCTGAAGAAAGAAATGGAATGAGAACATTTCA 3179  
Qy 3061 AGTACAGTGAACAAATTTAGCCGTAAATACATTAAGAAAAATGTTTAAAGAACCCAGC 3120  
Dd 3180 AGTACAGTGAACAAATTTAGCCGTAAATACATTAAGAAAAATGTTTAAAGAACCCAGC 3239  
Qy 3121 TCAGCAATATTAATGAAGTACGTTCCAGTACTAATGAAGTGGGCTCAGATTAATGA 3180  
Dd 3240 TCAGCAATATTAATGAAGTACGTTCCAGTACTAATGAAGTGGGCTCAGATTAATGA 3239  
Qy 3181 ATAGGTTCCAGTGAATGAACAAATTCAGCAGAACTAGGTAGAAGACAGAGGCCAAATTG 3240  
Dd 3300 ATAGGTTCCAGTGAATGAACAAATTCAGCAGAACTAGGTAGAAGACAGAGGCCAAATTG 3359  
Qy 3241 AATGCTATGCTTGAATTAAGGCTTTTGCACCTGAGGCTTATTAACAAAGTCTTCCGGA 3300  
Dd 3360 AATGCTATGCTTGAATTAAGGCTTTTGCACCTGAGGCTTATTAACAAAGTCTTCCGGA 3419  
Qy 3301 AGTAAATGTAACATCTGTAATTAAGAAAGCAAGAAATGGAAGAGTTCAGACTGCTT 3360  
Dd 3420 AGTAAATGTAACATCTGTAATTAAGAAAGCAAGAAATGGAAGAGTTCAGACTGCTT 3479  
Qy 3361 AATACAGATTTCTCTCATATCTGATTTTCAGATTAACCTTAGAACACCTTAGGGAAGT 3420  
Dd 3480 AATACAGATTTCTCTCATATCTGATTTTCAGATTAACCTTAGAACACCTTAGGGAAGT 3539  
Qy 3421 CATGATCTCAGGTTGTTCTGAGACCTGATGACCTGATAGATGATGGAATTAAG 3480  
Dd 3540 CATGATCTCAGGTTGTTCTGAGACCTGATGACCTGATAGATGATGGAATTAAG 3539  
Qy 3481 GAAGATACTAGTTTGTCTGCAAAATGACATTAAGAAAGTTCGCTTTTACCAAAAGC 3540  
Dd 3600 GAAGATACTAGTTTGTCTGCAAAATGACATTAAGAAAGTTCGCTTTTACCAAAAGC 3659  
Qy 3541 GTCCAGAAAGAGAGCTTACAGAGAGCTGACCTTTACCCATACATTTGGCTCAG 3600  
Dd 3660 GTCCAGAAAGAGAGCTTACAGAGAGCTGACCTTTACCCATACATTTGGCTCAG 3719  
Qy 3601 GGTACCGAAGAGGGGCCAAGAAATTAAGAGCTCAGAGAGAACTTTCAGTGAAGAT 3660  
Dd 3720 GGTACCGAAGAGGGGCCAAGAAATTAAGAGCTCAGAGAGAACTTTCAGTGAAGAT 3779  
Qy 3661 GAAGAGCTTCCCTGCTTCCAAACATTTGTTTGTAAAGTAACAAATATCTTCTCAG 3720  
Dd 3780 GAAGAGCTTCCCTGCTTCCAAACATTTGTTTGTAAAGTAACAAATATCTTCTCAG 3839  
Qy 3721 TCTACTAGCATAGCAGCGTGTCTACGAGTGTCTGTAAAGAACACAGAGAAATTTA 3780  
Dd 3840 TCTACTAGCATAGCAGCGTGTCTACGAGTGTCTGTAAAGAACACAGAGAAATTTA 3839  
Qy 3781 TTATCATTAAGAAATAGCTTAATGATGCTGACGTAAACCGTAATATTGGCAAGGCACT 3840  
Dd 3900 TTATCATTAAGAAATAGCTTAATGATGCTGACGTAAACCGTAATATTGGCAAGGCACT 3959  
Qy 3841 CAGAACATCACCTTAGTGAAGAAACAAATGTTCTGCTAGCTTGTCTTCCACAGTGC 3900  
Dd 3960 CAGAACATCACCTTAGTGAAGAAACAAATGTTCTGCTAGCTTGTCTTCCACAGTGC 4019  
Qy 3901 ACTGAATTTGAAGACTTGAATCAATTAACAAACCCAGAGATCTTCTTCTGATTTGTTCT 3960  
Dd 4020 ACTGAATTTGAAGACTTGAATCAATTAACAAACCCAGAGATCTTCTTCTGATTTGTTCT 4079  
Qy 3961 TCCAAACAAATGAGCATCAGTCTGAAGCCAGGAGTGTGCTGAGGACAAAGAAATG 4020  
Dd 4080 TCCAAACAAATGAGCATCAGTCTGAAGCCAGGAGTGTGCTGAGGACAAAGAAATG 4139

QY 4021 GTTTCAGATGATGAAGAAAGAGAACGCGCTTGGAGAAATAATACAGAGACAAAGC 4080  
 Db 4140 GTTTTCAGATGATGAAGAAAGAGAACGCGCTTGGAGAAATAATACAGAGACAAAGC 4199  
 QY 4081 ATGAGTCAAACTTAGTGAAGAGCATCTGGGTGAGAGTGAACAAGCCGCTCTGAA 4140  
 Db 4200 ATGAGTCAAACTTAGTGAAGAGCATCTGGGTGAGAGTGAACAAGCCGCTCTGAA 4259  
 QY 4141 GACGTGCAAGGGCTATCTCTCAGAGTACATTTTAAACCACTCAGAGAGGATACCATG 4200  
 Db 4260 GACTGCTCAGGGGTATCTCTCAGAGTACATTTTAAACCACTCAGAGAGGATACCATG 4319  
 QY 4201 CAACATTAACCTGTATTAAGCTCCAGAGAAATGGCTGAATAGACCTGTGTTAGAACG 4260  
 Db 4320 CAACATTAACCTGTATTAAGCTCCAGAGAAATGGCTGAATAGACCTGTGTTAGAACG 4379  
 QY 4261 CATGGAGGCAAGCCTTCTAAGACATACCTTCCATATAAGTACCTCTTGGCCCTTGAG 4320  
 Db 4380 CATGGAGGCAAGCCTTCTAAGACATACCTTCCATATAAGTACCTCTTGGCCCTTGAG 4439  
 QY 4321 GACCTGCAAAATCCAGAACAAACACATCAGAAAAGAGTATTAACCTTACAGAAAGT 4380  
 Db 4440 GACCTGCAAAATCCAGAACAAACACATCAGAAAAGAGTATTAACCTTACAGAAAGT 4499  
 QY 4381 AGTGAATACCTGTATTAAGCTCCAGAGAAATGGCTGTGCTGACAAAGTTGAGGTGCT 4440  
 Db 4500 AGTGAATACCTGTATTAAGCTCCAGAGAAATGGCTGTGCTGACAAAGTTGAGGTGCT 4559  
 QY 4441 GCAGATAGTTCATCCAGTAAATTAAGAACCCAGAGGTGAAGAGTTCATCCCTTTAAA 4500  
 Db 4560 GCAGATAGTTCATCCAGTAAATTAAGAACCCAGAGGTGAAGAGTTCATCCCTTTAAA 4619  
 QY 4501 TGCCATCATTTAGATAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 4560  
 Db 4620 TGCCATCATTTAGATAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 4679  
 QY 4561 AACTACCATCTCAGAGAGAGCTCATTAAGTTGTTGATGATGATGATGATGATGATGATG 4620  
 Db 4680 AACTACCATCTCAGAGAGAGCTCATTAAGTTGTTGATGATGATGATGATGATGATGATG 4739  
 QY 4621 GAGTGGGGCCACAGATTTGAGCGAAACATCTTACTTGGCAGGAAATCTAAGAGGA 4680  
 Db 4740 GAGTGGGGCCACAGATTTGAGCGAAACATCTTACTTGGCAGGAAATCTAAGAGGA 4799  
 QY 4681 ACCCCTTACCTGGAATCTGGAATCAGCCTTCTCTGATGACCTGATCTGATCTTCT 4740  
 Db 4800 ACCCCTTACCTGGAATCTGGAATCAGCCTTCTCTGATGACCTGATCTGATCTTCT 4859  
 QY 4741 GAAGACAGAGCCCGAGAGTCAAGCTGTGTTGGCAACATACCATCTTCAACCTCTGCA 4800  
 Db 4860 GAAGACAGAGCCCGAGAGTCAAGCTGTGTTGGCAACATACCATCTTCAACCTCTGCA 4919  
 QY 4801 AAAGTTCCTCCCAATGAAAGTTGAGAAATGAGCAATCTTACTTCCAGAGGCAATCTA 4860  
 Db 4920 AAAGTTCCTCCCAATGAAAGTTGAGAAATGAGCAATCTTACTTCCAGAGGCAATCTA 4979  
 QY 4861 GATACTGCTGGGTATTAATGCAATGAGAAAGTGTGAGCAGGAGAGAGCAGAAATGACA 4920  
 Db 4980 GATACTGCTGGGTATTAATGCAATGAGAAAGTGTGAGCAGGAGAGAGCAGAAATGACA 5039  
 QY 4921 GCTTCAACAGAAAGGCTCAACAAAGATGTCATGTTGCTGCTGAGCCCGAGAA 4980  
 Db 5040 GCTTCAACAGAAAGGCTCAACAAAGATGTCATGTTGCTGCTGAGCCCGAGAA 5099  
 QY 4981 GAATTTATCTGCTGTACAAAGTTTGGCAAAAACACACATCACTTAACTAATCTAAT 5040  
 Db 5100 GAATTTATCTGCTGTGTACAAAGTTTGGCAAAAACACACATCACTTAACTAATCTAAT 5159  
 QY 5041 ACTGAAGAGTACTATCTGTTTATGAAGACAGATGCTGAGTTGTGTGAAGGAGACA 5100  
 Db 5160 ACTGAAGAGTACTATCTGTTTATGAAGACAGATGCTGAGTTGTGTGAAGGAGACA 5219  
 QY 5101 CTGAATATTTTCTAGGAATTCGGGAGGAGAAATGGGTAGTTAGCTATTTCTGGGTGACC 5160

Db 5220 CTGAATATTTTCTAGGAATTCGGGAGGAGAAATGGGTAGTTAGCTATTTCTGGGTGACC 5279  
 QY 5161 CAGCTATTTAAAGAAAGAAATGCTGAATGAGATGATTTTGAAGCAGAGAGATGTC 5220  
 Db 5280 CAGCTATTTAAAGAAAGAAATGCTGAATGAGATGATTTTGAAGCAGAGAGATGTC 5339  
 QY 5221 GTCAATGAGAAAGAACCAAGGTCGCAAGAGCAGAGAAATCCAGAGCAGAAAGATC 5280  
 Db 5340 GTCAATGAGAAAGAACCAAGGTCGCAAGAGCAGAGAAATCCAGAGCAGAAAGATC 5399  
 QY 5281 TTCAGGGGCTTGAATCTGTGCTATGAGGCTTTCACCAACATGCCCAGATCAACTG 5340  
 Db 5400 TTCAGGGGCTTGAATCTGTGCTATGAGGCTTTCACCAACATGCCCAGATCAACTG 5459  
 QY 5341 GAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5400  
 Db 5460 GAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5519  
 QY 5401 GGCACAGGTGTCACCAATTTGTTGTTGTCAGACCCAGATGCTGACAGAGCAATGGC 5460  
 Db 5520 GGCACAGGTGTCACCAATTTGTTGTTGTCAGACCCAGATGCTGACAGAGCAATGGC 5579  
 QY 5461 TTCATGCAATTTGGGAGATGTTGAGGCACTGTGTTGACCCGAGAGTGGTGTGGAC 5520  
 Db 5580 TTCATGCAATTTGGGAGATGTTGAGGCACTGTGTTGACCCGAGAGTGGTGTGGAC 5639  
 QY 5521 AGTGTAGCACTCTACCAAGTCCAGAGAGCTGAGACACTTACTGATACCCAGATCCCC 5580  
 Db 5640 AGTGTAGCACTCTACCAAGTCCAGAGAGCTGAGACACTTACTGATACCCAGATCCCC 5699  
 QY 5581 AGCCACTAC 5589  
 Db 5700 AGCCACTAC 5708

RESULT 3  
 US-09-734-672-5  
 : Sequence 5, Application US/09734672  
 : Publication No. US20020183268A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Murphy, Patricia D.  
 : Allen, Antonette C.  
 : Alvarez, Christopher P.  
 : Critz, Brenda S.  
 : Olson, Sheri J.  
 : Schellier, Denise B.  
 : Zeng, Bin  
 : TITLE OF INVENTION: Coding Sequences of the Human  
 : BRCA1 Gene  
 : NUMBER OF SEQUENCES: 72  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Morgan Lewis & Bockius LLP  
 : STREET: 1111 Pennsylvania Ave., N.W.  
 : CITY: Washington  
 : STATE: District of Columbia  
 : COUNTRY: USA  
 : ZIP: 20004  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: PatentIn Release #1.0, Version #1.30  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/09/734,672  
 : FILING DATE: 03-Dec-2000  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: US 08/966,436  
 : FILING DATE: 07-No. US20020183268A1-97  
 : APPLICATION NUMBER: US 08/598,591  
 : FILING DATE: 12-Feb-96  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: Michael S. Tuscan

REGISTRATION NUMBER: 43,210  
REFERENCE/DOCK NUMBER: 44921-5055-02-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-739-3000  
TELEFAX: 202-739-3001  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5711 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: No. US20020183268A1 Relevant  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
STRAIN: BRCAL  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: 17  
MAP POSITION: 17q21  
SEQUENCE DESCRIPTION: SEQ ID NO: 5:  
US-09-734-672-5

Query Match 99.9% Score 5582.6 DB 9 Length 5711:

Best Local Similarity 99.9% Pred. No. 0:  
Matches 5585: Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
QY 1 ATGATTTATCTGCTTCGCGCTTGAAGAGTACAAATGCTATATGCTATGCAGAAA 60
DB 120 ATGATTTATCTGCTTCGCGCTTGAAGAGTACAAATGCTATATGCTATGCAGAAA 179
QY 61 ATCTTAGAGTCCCATCTGCTGAGTGTATCAGAGAACTGTCTCCCAAAAGTGTAC 120
DB 180 ATCTTAGAGTCCCATCTGCTGAGTGTATCAGAGAACTGTCTCCCAAAAGTGTAC 239
QY 121 CACATATTTTGAAGATTTTGCATGCGAATCTTCAACAGAGAAAGGCGCTTACAG 180
DB 240 CACATATTTTGAAGATTTTGCATGCGAATCTTCAACAGAGAAAGGCGCTTACAG 299
QY 181 TGTCTTATGTAGAGATGATATTAACCAAGAGAGCTTCAAGAAAGTACGAGATTAGT 240
DB 300 TGTCTTATGTAGAGATGATATTAACCAAGAGAGCTTCAAGAAAGTACGAGATTAGT 359
QY 241 CAAGTGTGTAAGAGTATGAAATCATTTGCTTTCAGCTTACAGTACAGTTGGAG 300
DB 360 CAAGTGTGTAAGAGTATGAAATCATTTGCTTTCAGCTTACAGTACAGTTGGAG 419
QY 301 TATGCAACAGCTATATATTTTGCAGAAAGGAAATTAACCTCTCGAATCATTAAGAT 360
DB 420 TATGCAACAGCTATATATTTTGCAGAAAGGAAATTAACCTCTCGAATCATTAAGAT 479
QY 361 GAAGTTCTATATCAATCAAGATATGAGCTACAGAAACCGTGCAGAAAGACTTCTACAG 420
DB 480 GAAGTTCTATATCAATCAAGATATGAGCTACAGAAACCGTGCAGAAAGACTTCTACAG 539
QY 421 GAACCCGAAATTCCTCTGAGAGAAACCACTCTCAGTGTCAACTCTCTAACCTTGA 480
DB 540 GAACCCGAAATTCCTCTGAGAGAAACCACTCTCAGTGTCAACTCTCTAACCTTGA 599
QY 481 ACTGTGAGAGCTCTGAGAGCAAGAGAGGATACCACTCAAAAGAGCGTCTGTACAT 540
DB 600 ACTGTGAGAGCTCTGAGAGCAAGAGAGGATACCACTCAAAAGAGCGTCTGTACAT 659
QY 541 GAATTTGGATCTGATTTCTGAGAGATACCGTTAATTAAGCAACTTATTCAGTGTGGA 600
DB 660 GAATTTGGATCTGATTTCTGAGAGATACCGTTAATTAAGCAACTTATTCAGTGTGGA 719
QY 601 GATCAGAGATTTGTTCAAAATCAACCCCTCAAGAGAACAGGATGAATCAGTTTGATCT 660
DB 720 GATCAGAGATTTGTTCAAAATCAACCCCTCAAGAGAACAGGATGAATCAGTTTGATCT 779
QY 661 GCAAAAAAGGCTGCTGTGAATTTCTGAGAGGATGTAACAAATACGAAATCATATCAA 720
DB 780 GCAAAAAAGGCTGCTGTGAATTTCTGAGAGGATGTAACAAATACGAAATCATATCAA 839
```

```
QY 721 CCCAGTAATTAATGATTTTGAACACACTGAGAGCGTGCAGTGTAGAGCATCCAGAAAG 780
DB 840 CCCAGTAATTAATGATTTTGAACACACTGAGAGCGTGCAGTGTAGAGCATCCAGAAAG 839
QY 781 TATCAGAGTATGTTCTTTTCAAACTTGCATGTGAGGCAATGTGACAAATATCTATCC 840
DB 900 TATCAGAGTATGTTCTTTTCAAACTTGCATGTGAGGCAATGTGACAAATATCTATCC 959
QY 841 AGCTCATTCAGCATGAGAGACAGCACTTATTAATGATGAGGAGGAGGAGGAGGAGG 900
DB 960 AGCTCATTCAGCATGAGAGACAGCACTTATTAATGATGAGGAGGAGGAGGAGGAGG 1019
QY 901 AAGGCTGAATTCGTATATAAAGCAACAGCTGCTTACAGAGAGGCAACATATACGA 960
DB 1020 AAGGCTGAATTCGTATATAAAGCAACAGCTGCTTACAGAGAGGCAACATATACGA 1079
QY 961 TGGGCTGGAAGTAAAGAAACATGTAAATGATGAGGAGCTCCAGACAGAAAAAGGTA 1020
DB 1080 TGGGCTGGAAGTAAAGAAACATGTAAATGATGAGGAGCTCCAGACAGAAAAAGGTA 1139
QY 1021 GATCTGAATGCTGATCCCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
DB 1140 GATCTGAATGCTGATCCCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1199
QY 1081 TCAGAGAAATCTAGAGATATCTGAAGATGTCCTTGATTAACATTAATATGAGCATTCAG 1140
DB 1200 TCAGAGAAATCTAGAGATATCTGAAGATGTCCTTGATTAACATTAATATGAGCATTCAG 1259
QY 1141 AAGGTTAATGAGTGGTTTCCAGAAAGTATGAATCTTGAATGCTGATGACTCATGAT 1200
DB 1260 AAGGTTAATGAGTGGTTTCCAGAAAGTATGAATCTTGAATGCTGATGACTCATGAT 1319
QY 1201 GGGAGCTGATTAATCAAAATCCCAAGTACGTATGATGAGAGCTTAAATGAGGTTGAT 1260
DB 1320 GGGAGCTGATTAATCAAAATCCCAAGTACGTATGATGAGAGCTTAAATGAGGTTGAT 1379
QY 1261 GAATATTCGTTCTTCAGAGAAATAGAGTACAGGAGGAGTATGATGATGAGGCTTTA 1320
DB 1380 GAATATTCGTTCTTCAGAGAAATAGAGTACAGGAGGAGTATGATGATGAGGCTTTA 1439
QY 1321 ATATGTAAGAGTGAAGAGTTCATCTCAATCAGTAGAGAGTAAATTTGAAGACAAATA 1380
DB 1440 ATATGTAAGAGTGAAGAGTTCATCTCAATCAGTAGAGAGTAAATTTGAAGACAAATA 1499
QY 1381 TTTGGGAAACCTATTCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440
DB 1500 TTTGGGAAACCTATTCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1559
QY 1441 CTAATTTATGAGAGATTTTCTACTGAGCAGAGATTAATACAGAGCGTCCCTCACAAAT 1500
DB 1560 CTAATTTATGAGAGATTTTCTACTGAGCAGAGATTAATACAGAGCGTCCCTCACAAAT 1619
QY 1501 AAATTTAAGCGTAAAGAGAGCTTACATCAGGCGCTTCAATCTGAGAGATTTTATCAAGAA 1560
DB 1620 AAATTTAAGCGTAAAGAGAGCTTACATCAGGCGCTTCAATCTGAGAGATTTTATCAAGAA 1679
QY 1561 GCAGATTTGGAGCTTCAAAAGAGCTCTGAATTAATTCAGGAGACTTAACCAAGGAG 1620
DB 1680 GCAGATTTGGAGCTTCAAAAGAGCTCTGAATTAATTCAGGAGACTTAACCAAGGAG 1739
QY 1621 CAGATGCTCAGTGAATATTAATTAATAGTGTGATGAGATTAATTAACAAAGGATGAT 1680
DB 1740 CAGATGCTCAGTGAATATTAATTAATAGTGTGATGAGATTAATTAACAAAGGATGAT 1799
QY 1681 TCTATTTCAGATGAGAGAGAGCTTACCAATACAAATCAGTCAAAAAAGATCTGCTTTC 1740
DB 1800 TCTATTTCAGATGAGAGAGAGCTTACCAATACAAATCAGTCAAAAAAGATCTGCTTTC 1859
QY 1741 AAAAGGAAAGCTGAACCTATTAAGAGAGAGATTAAGCAATTAAGCAATTAAGCAAT 1800
DB 1860 AAAAGGAAAGCTGAACCTATTAAGAGAGAGATTAAGCAATTAAGCAATTAAGCAAT 1919
QY 1801 CACAATTCAAAAGACCTTAAGAGAGAGATTAAGCAATTAAGCAATTAAGCAATTAAT 1860
```

|||||  
Db 1920 CACATTTCAAAAGACCTTAATAAAGAGTGTGAGGAGGAGTCTTCTACAGGACAAATTT 1979  
Qy 1861 CATGCGCTTAAGACTAGTACAGTAAATCTAAGCCACCTAATTTAGTGAATTCGAA 1920  
Db 1980 CATGCGCTTAAGACTAGTACAGTAAATCTAAGCCACCTAATTTAGTGAATTCGAA 2039  
Qy 1921 ATTGATAGTGTCTTACAGTGAAGAGATAAAGAAAAAGTACAAACCAATGCCAGTC 1980  
Db 2040 ATTGATAGTGTCTTACAGTGAAGAGATAAAGAAAAAGTACAAACCAATGCCAGTC 2099  
Qy 1981 AGGACACGACGAAACCTTACAACTCATGGAAGTTAAAGAACTGCACTGGAGCCAGAG 2040  
Db 2100 AGGACACGACGAAACCTTACAACTCATGGAAGTTAAAGAACTGCACTGGAGCCAGAG 2159  
Qy 2041 AGTAAACAGCCAAATGAAAGACAGAAAGTAAAGACATGACATGATCTTCCAGAGCTG 2100  
Db 2160 AGTAAACAGCCAAATGAAAGACAGAAAGTAAAGACATGACATGATCTTCCAGAGCTG 2219  
Qy 2101 AAGTTAAACAATGACACCTGTTCTTTACTAAGTGTTCAAATACAGTGAACCTTAAAGAA 2160  
Db 2220 AAGTTAAACAATGACACCTGTTCTTTACTAAGTGTTCAAATACAGTGAACCTTAAAGAA 2279  
Qy 2161 TTTGTCAATCCTTACCTTTCCAGAGAAAGAAAGAGAACTAGAAACAGTTAAAGTG 2220  
Db 2280 TTTGTCAATCCTTACCTTTCCAGAGAAAGAAAGAGAACTAGAAACAGTTAAAGTG 2339  
Qy 2221 TCTAATATGCTGAAGACCCCAAGATCTCATGTTAAGTGGAGAAAGGTTTCCAACT 2280  
Db 2340 TCTAATATGCTGAAGACCCCAAGATCTCATGTTAAGTGGAGAAAGGTTTCCAACT 2399  
Qy 2281 GAAAGATCTGTAGAGTAGAGATATTTTCATTTGTTAGTGTGTTATGAGCAGTCTAG 2340  
Db 2400 GAAAGATCTGTAGAGTAGAGATATTTTCATTTGTTAGTGTGTTATGAGCAGTCTAG 2459  
Qy 2341 GAAAGATCTGTAGAGTAGAGATATTTTCATTTGTTAGTGTGTTATGAGCAGTCTAG 2400  
Db 2460 GAAAGATCTGTAGAGTAGAGATATTTTCATTTGTTAGTGTGTTATGAGCAGTCTAG 2519  
Qy 2401 TGTGTAGTGTGAGTGTGAGATTTTGAAGACCCCAAGGAGTAAATTCATGTTTCCAAA 2460  
Db 2520 TGTGTAGTGTGAGTGTGAGATTTTGAAGACCCCAAGGAGTAAATTCATGTTTCCAAA 2579  
Qy 2461 GATTAATGAAATGACACAGAGAGCTTTAAGTATTCATTTGGACATGAAGTTAACACAGT 2520  
Db 2580 GATTAATGAAATGACACAGAGAGCTTTAAGTATTCATTTGGACATGAAGTTAACACAGT 2639  
Qy 2521 CGGGAACAAGCATGAAATGGAAGAAAGTGAATCTGCTCATGTTTGCAGAAATACA 2580  
Db 2640 CGGGAACAAGCATGAAATGGAAGAAAGTGAATCTGCTCATGTTTGCAGAAATACA 2699  
Qy 2581 TTCAAGGTTTCAAAAGCCGACATTTGCTGTGTTTCAATCCAGGAAATGCGAAGAG 2640  
Db 2700 TTCAAGGTTTCAAAAGCCGACATTTGCTGTGTTTCAATCCAGGAAATGCGAAGAG 2759  
Qy 2641 GAATGTGCAACATTTCTGCCCACCTCTGGTCTTTAAAGAAACAAGTCCAAAAGTCACT 2700  
Db 2760 GAATGTGCAACATTTCTGCCCACCTCTGGTCTTTAAAGAAACAAGTCCAAAAGTCACT 2819  
Qy 2701 TTTGATGTGCAAAAGAAAGAAATCAAGAAAGATGAGTCAATATCAACGCTGTA 2760  
Db 2820 TTTGATGTGCAAAAGAAAGAAATCAAGAAAGATGAGTCAATATCAACGCTGTA 2879  
Qy 2761 CAGACATTAATATCACTGACAGGCTTCTGTTGTGTCAGAAAGATAGCCAGTGTAT 2820  
Db 2880 CAGACATTAATATCACTGACAGGCTTCTGTTGTGTCAGAAAGATAGCCAGTGTAT 2939  
Qy 2821 AATGCCAATATGATATCAAGAGAGCTAGGTTTGTATCATCTCAAGTTCAGAGGC 2880  
Db 2940 AATGCCAATATGATATCAAGAGAGCTAGGTTTGTATCATCTCAAGTTCAGAGGC 2999  
Qy 2881 AAGCAAACTGAGCTCATTTCTCCAAATTAACATGAGCTTTTACAAAACCATATCGTATA 2940  
|||||

Db 3000 AACGAACCTGAGCTCATTTACTCCAAATTAACATGAGCTTTTACAAAACCCATATCGTATA 3059  
Qy 2941 CCACCACCTTTTCCCATCAAGTCAATTTGTAAATCAATATGAGAAAAATCTGCTAGAG 3000  
Db 3060 CCACCACCTTTTCCCATCAAGTCAATTTGTAAATCAATATGAGAAAAATCTGCTAGAG 3119  
Qy 3001 GAAACCTTTGAGGAACATTCATGTCACCTGAAGAGAAATGGAATGAGACATTTCCA 3060  
Db 3120 GAAACCTTTGAGGAACATTCATGTCACCTGAAGAGAAATGGAATGAGACATTTCCA 3179  
Qy 3061 AGTACAGTGAACAAATTTAGCCGTAATATACCTTTGAGAAATGTTTTAAAGAACCCAGC 3120  
Db 3180 AGTACAGTGAACAAATTTAGCCGTAATATACCTTTGAGAAATGTTTTAAAGAACCCAGC 3239  
Qy 3121 TCAAGCAATATTAATGAAGTAGTGTCCAGTCTATGTAAGTGGGCTCCAGTATTAATGAA 3180  
Db 3240 TCAAGCAATATTAATGAAGTAGTGTCCAGTCTATGTAAGTGGGCTCCAGTATTAATGAA 3299  
Qy 3181 ATAGGTTCCAGTGTGAAACATTTCAAGCAGAACTAGTAGAAAAGAGGCCAAATATG 3240  
Db 3300 ATAGGTTCCAGTGTGAAACATTTCAAGCAGAACTAGTAGAAAAGAGGCCAAATATG 3359  
Qy 3241 AATGCTATGCTTACATTAAGGGGTTTGTGCAACCTGAGGCTTTAAACAAGTCTTCTGGA 3300  
Db 3360 AATGCTATGCTTACATTAAGGGGTTTGTGCAACCTGAGGCTTTAAACAAGTCTTCTGGA 3419  
Qy 3301 AGTAAATGTAAGCATCTCGAATAAAGAAAGCAAGAAATGTAAGAAAGTACAGACGTT 3360  
Db 3420 AGTAAATGTAAGCATCTCGAATAAAGAAAGCAAGAAATGTAAGAAAGTACAGACGTT 3479  
Qy 3361 AATACAGATTTCTCTCCATATCTGATTTACAGATTAATGTAAGAAAGCTTATGGAATAGT 3420  
Db 3480 AATACAGATTTCTCTCCATATCTGATTTACAGATTAATGTAAGAAAGCTTATGGAATAGT 3539  
Qy 3421 CATGATCTCAGGTTTGTGTCGAGACACCTGATGACCTGTTAGTGTGTAAGTAAG 3480  
Db 3540 CATGATCTCAGGTTTGTGTCGAGACACCTGATGACCTGTTAGTGTGTAAGTAAG 3599  
Qy 3481 GAAGATACTAGTTTGTGGAATAAGTACATTAAGAAAGTGTGTTTAAAGCAAAAGC 3540  
Db 3600 GAAGATACTAGTTTGTGGAATAAGTACATTAAGAAAGTGTGTTTAAAGCAAAAGC 3659  
Qy 3541 GTCCAGAAAGAGAGCTTAGCAGAGTCTTAGCCCTTACCCATACATTTGGCTCAG 3600  
Db 3660 GTCCAGAGAGAGAGCTTAGCAGAGTCTTAGCCCTTACCCATACATTTGGCTCAG 3719  
Qy 3601 GGTATCCGAAGAGGGGCCAAGAAATTAAGATGCTCAGAAAGAACTTATCTAGAGAGAT 3660  
Db 3720 GGTATCCGAAGAGGGGCCAAGAAATTAAGATGCTCAGAAAGAACTTATCTAGAGAGAT 3779  
Qy 3661 GAAGAGCTTCCCTGCTTCCAACTGTTGTTATTTGTTAAGTAAACATATATCTTCTCAG 3720  
Db 3780 GAAGAGCTTCCCTGCTTCCAACTGTTGTTATTTGTTAAGTAAACATATATCTTCTCAG 3839  
Qy 3721 TCTACTAGGCAATACACCGTTTGTCTACAGAGTGTCTGTTAAGAACACAGAGAGAAATTTA 3780  
Db 3840 TCTACTAGGCAATACACCGTTTGTCTACAGAGTGTCTGTTAAGAACACAGAGAGAAATTTA 3899  
Qy 3781 TTATCATTTGAAGATTAAGTATGATGCACTGCACTAAGCAAGTAAATTTGGCAAGCATCT 3840  
Db 3900 TTATCATTTGAAGATTAAGTATGATGCACTGCACTAAGCAAGTAAATTTGGCAAGCATCT 3959  
Qy 3841 CAGGAACATCACTTATGAGGAAACAAAGTGTCTGCTGCTGTTTCTTCCAGAGTC 3900  
Db 3960 CAGGAACATCACTTATGAGGAAACAAAGTGTCTGCTGCTGTTTCTTCCAGAGTC 4019  
Qy 3901 AGTGAATTTGGAAGACTTACCTGCAAAATACAAACACCCAGAGATCTTCTTCTGATGTTCT 3960  
Db 4020 AGTGAATTTGGAAGACTTACCTGCAAAATACAAACACCCAGAGATCTTCTTCTGATGTTCT 4079  
Qy 3961 TCCAAACAATGAGGCACTCACTGTGAAAGCCAGGAGTGTGCTGAGTGAAGAAATG 4020  
Db 4080 TCCAAACAATGAGGCACTCACTGTGAAAGCCAGGAGTGTGCTGAGTGAAGAAATG 4139  
|||||

4021 GTTTCAGATGATGAAGAAGAGGAGCGGCTTGGAAAGAAATATATCAGAAGCAAGC 4080  
|||||  
4140 GTTTCAGATGATGAAGAAGAGGAGCGGCTTGGAAAGAAATATATCAGAAGCAAGC 4139  
4081 ATGAGATTCAAACTTAGGTGAAGCAGCATCTGGGTGTGAGAGTGAACAAAGCGTCTGAA 4140  
|||||  
4200 ATGAGATTCAAACTTAGGTGAAGCAGCATCTGGGTGTGAGAGTGAACAAAGCGTCTGAA 4259  
4141 GACTGCTCAGGCGTATCTCTCAGAGTGAATTTTAACCACTCAGCAGAGGAGTACATG 4200  
|||||  
4260 GACTGCTCAGGCGTATCTCTCAGAGTGAATTTTAACCACTCAGCAGAGGAGTACATG 4319  
4201 CAACATTAACCTGATTAAGCTCCAGCAGAAATGCTGAATAGAGCTGTGTAAGACAG 4260  
|||||  
4320 CAACATTAACCTGATTAAGCTCCAGCAGAAATGCTGAATAGAGCTGTGTAAGACAG 4379  
4261 CATGGAGGCGAGCGCTTCTAAGCAGCTACCTTCCATCATTAAGTACCTCTTCCCTTGAG 4320  
|||||  
4380 CATGGAGGCGAGCGCTTCTAAGCAGCTACCTTCCATCATTAAGTACCTCTTCCCTTGAG 4439  
4321 GACCTGCGCAATTCACAAACAGCAGATCGAAAAAGCAATTAATTAATTAATTAATTA 4380  
|||||  
4440 GACCTGCGCAATTCACAAACAGCAGATCGAAAAAGCAATTAATTAATTAATTAATTA 4499  
4381 AGTGAATACCTCTAATAGCAGATCCAGAGGCGCTTCTGCTGACAAAGTTGAGGTGCT 4440  
|||||  
4500 AGTGAATACCTCTAATAGCAGATCCAGAGGCGCTTCTGCTGACAAAGTTGAGGTGCT 4559  
4441 GCAGATAGTTCTACCAAGTAAAAATTAAGAACCAAGAGTGAAGAGTCAATCCCTTCTAA 4500  
|||||  
4560 GCAGATAGTTCTACCAAGTAAAAATTAAGAACCAAGAGTGAAGAGTCAATCCCTTCTAA 4619  
4501 TGCCCATCATTTAGATAGTGTGTACATGACAGATGCTCTGGAGTCTTCAGAAATAGA 4560  
|||||  
4620 TGCCCATCATTTAGATAGTGTGTACATGACAGATGCTCTGGAGTCTTCAGAAATAGA 4679  
4561 AACTACCATCTCAAGAGGAGCTCATTAAGTGTGTGATGTGAGAGCAACAGCTGAA 4620  
|||||  
4680 AACTACCATCTCAAGAGGAGCTCATTAAGTGTGTGATGTGAGAGCAACAGCTGAA 4739  
4621 GAGTGTGGGCGCACAGATTTGACGGAAACATCTTACTTGCACAGCAAGATCTAGAGGA 4680  
|||||  
4740 GAGTGTGGGCGCACAGATTTGACGGAAACATCTTACTTGCACAGCAAGATCTAGAGGA 4799  
4681 ACCCCTTACCTGGAATCTGGAATCAGCTCTTCTGATGAGCCCTGAATCTGATCTCT 4740  
|||||  
4800 ACCCCTTACCTGGAATCTGGAATCAGCTCTTCTGATGAGCCCTGAATCTGATCTCTCT 4859  
4741 GAAGACAGAGCCCGACAGTCACTGCTGTGGCAACATACCATCTTCAACCTCTGATG 4800  
|||||  
4860 GAAGACAGAGCCCGACAGTCACTGCTGTGGCAACATACCATCTTCAACCTCTGATG 4919  
4801 AAAGTTCCCAATTTGAAGTTGACAGATCTGCCAGAGTCCAGCTGCTCTACTACT 4860  
|||||  
4920 AAAGTTCCCAATTTGAAGTTGACAGATCTGCCAGAGTCCAGCTGCTCTACTACT 4979  
4861 GATACGCTGGGTATTAATGCAATGGAAGAGTGTGAGAGGAGGAGGAGGAGGAGGAG 4920  
|||||  
4980 GATACGCTGGGTATTAATGCAATGGAAGAGTGTGAGAGGAGGAGGAGGAGGAGGAG 5039  
4921 GCTTCACAGAAAGGCTCAACAAAGAAATGCTCATGCTGTGCTGCTGACCCAGAA 4980  
|||||  
5040 GCTTCACAGAAAGGCTCAACAAAGAAATGCTCATGCTGTGCTGCTGACCCAGAA 5099  
4981 GAATTTATGCTGTGTGACAGTTTGCAGAAACACACATCACTTTAACTAATCTAAT 5040  
|||||  
5100 GAATTTATGCTGTGTGACAGTTTGCAGAAACACACATCACTTTAACTAATCTAAT 5159  
5041 ACTGAAGAGTACTACATGTTGTATGAAGAAACAGATGAGTGTGTGTGAGAGGAG 5100  
|||||  
5160 ACTGAAGAGTACTACATGTTGTATGAAGAAACAGATGAGTGTGTGTGAGAGGAG 5219

QY 5101 CTGAATATTTTCTAGGAATTCGGGAGGAAAAATGGGTAGTAAAGTATTTCTGGGTGAC 5160  
|||||  
Db 5220 CTGAATATTTTCTAGGAATTCGGGAGGAAAAATGGGTAGTAAAGTATTTCTGGGTGAC 5279  
QY 5161 CAGCTATTTAAGAAAGAAAAATGCTGAATGACATGATTTTGAAGTCAGAGAGATGTG 5220  
|||||  
Db 5280 CAGCTATTTAAGAAAGAAAAATGCTGAATGACATGATTTTGAAGTCAGAGAGATGTG 5339  
QY 5221 GTCATGGAAGAAACCAACCAAGGTCCAAAGCCGAGCAAGAGATGCCAGACGAAGATC 5280  
|||||  
Db 5340 GTCATGGAAGAAACCAACCAAGGTCCAAAGCCGAGCAAGAGATGCCAGACGAAGATC 5399  
QY 5281 TTCAGGGGCTTACAAATCTGTCTATGGGCCCTTACCAAAATGCCCCAGATCAATG 5340  
|||||  
Db 5400 TTCAGGGGCTTACAAATCTGTCTATGGGCCCTTACCAAAATGCCCCAGATCAATG 5459  
QY 5341 GAATGATGTATACAGCTGTGTGTGCTTCTGTGTGTAAGAGCTTTTCATCATTCACCTT 5400  
|||||  
Db 5460 GAATGATGTATACAGCTGTGTGTGCTTCTGTGTGTAAGAGCTTTTCATCATTCACCTT 5519  
QY 5401 GGCACAGGTCTCCACCCTTGTGTGTGCTTGTGTGCAAGCCAGATGCTTGACAGACAAATGGC 5460  
|||||  
Db 5520 GGCACAGGTCTCCACCCTTGTGTGTGCTTGTGTGCAAGCCAGATGCTTGACAGACAAATGGC 5579  
QY 5461 TTCATGCAATTTGGCAGATGTGTGAGCAGCAGCTGTGTGACCCGAGAGTGGGTGTGAC 5520  
|||||  
Db 5580 TTCATGCAATTTGGCAGATGTGTGAGCAGCAGCTGTGTGACCCGAGAGTGGGTGTGAC 5639  
QY 5521 AGTGTAGCACTTACCAAGTGTGAGCAGAGCTGTGACACCTTACCTGATACCCAGATCCCCAC 5580  
|||||  
Db 5640 AGTGTAGCACTTACCAAGTGTGAGCAGAGCTGTGACACCTTACCTGATACCCAGATCCCCAC 5699  
QY 5581 AGCCACTAC 5589  
|||||  
Db 5700 AGCCACTAC 5708

RESULT 4  
US-09-982-828-3  
; Sequence 3, Application US/09982828  
; Publication No. US20030022184A1  
; GENERAL INFORMATION:  
; APPLICANT: Murphy, Patricia D.  
; Allen, Antonette C.  
; Alvarez, Christopher P.  
; Olson, Sheri J.  
; Thurber, Denise  
; Zeng, Bin  
; TITLE OF INVENTION: Coding Sequences of the Human  
; BRCAL Gene  
; NUMBER OF SEQUENCES: 72  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Morgan Lewis & Bockius LLP  
; STREET: 1111 Pennsylvania Avenue N. W.  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/982,828  
; FILING DATE: 22-Oct-2001  
; CLASSIFICATION: <unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/074,453  
; FILING DATE: 1998-05-06  
; APPLICATION NUMBER: US 08/798,691  
; FILING DATE: 1997-02-12



APPLICATION NUMBER: US 08/598,591  
FILING DATE: 1996-02-12  
ATTORNEY/AGENT INFORMATION:  
NAME: Michael S. Tuscan  
REGISTRATION NUMBER: 43,210  
REFERENCE/DOCKET NUMBER: 44921-5053-01-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-739-3001  
TELEFAX: 202-739-3000  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5711 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: not relevant  
TOPOLOGY: Linear  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
STRAIN: BRCA1 (cm12)  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: 17  
MAP POSITION: 17q21  
SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-09-982-828-3

Query Match 99.9%; Score 5582.6; DB 9; Length 5711;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 5585; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 ATGATATTATCTGCTCTCCGCTTGAAGAGTACAAAATGTCATTAACTATGACAGAAA 60  
DB 120 ATGATATTATCTGCTCTCCGCTTGAAGAGTACAAAATGTCATTAACTATGACAGAAA 179  
OY 61 ATCTAGAGTCCATCTGCTGAGTATGATCAAGAACCTGCTCCACAAGTGCAC 120  
DB 180 ATCTAGAGTCCATCTGCTGAGTATGATCAAGAACCTGCTCCACAAGTGCAC 239  
OY 121 CACATATTTTGAAGATTTTGCATGCTGAACCTTCTCAACAGAGAGGCGCTTCACAG 180  
DB 240 CACATATTTTGAAGATTTTGCATGCTGAACCTTCTCAACAGAGAGGCGCTTCACAG 299  
OY 181 TGTCCTTATGTAAGATATATACCAAGAGAGCTTACAGAAAGTACAGATTTAGT 240  
DB 300 TGTCCTTATGTAAGATATATACCAAGAGAGCTTACAGAAAGTACAGATTTAGT 359  
OY 241 CACTTGTGAAGAGATATGAAAATGATTCCTTTGACGTTGACAGAGTTGGAG 300  
DB 360 CACTTGTGAAGAGATATGAAAATGATTCCTTTGACGTTGACAGAGTTGGAG 419  
OY 301 TATGCAAAAGCTATATTTTGGCAAAAAGAAAATTAATCTCCTGTAACATCTTAAAGAT 360  
DB 420 TATGCAAAAGCTATATTTTGGCAAAAAGAAAATTAATCTCCTGTAACATCTTAAAGAT 479  
OY 361 GAAGTTTCTATCATCAAGATATGAGGCTACAGAAACCGTCCCAAGAGCTTCTACAGAGT 420  
DB 480 GAAGTTTCTATCATCAAGATATGAGGCTACAGAAACCGTCCCAAGAGCTTCTACAGAGT 539  
OY 421 GAACCCGAAAATCCTTCTGCAAGAAACAGCTGCTGTAACCTCTAACCCTTGA 480  
DB 540 GAACCCGAAAATCCTTCTGCAAGAAACAGCTGCTGTAACCTCTAACCCTTGA 599  
OY 481 ACTGTGAGAACCTGTGAGAGCAAGAGAGCTGTAACCTCTAACCCTTGA 540  
DB 600 ACTGTGAGAACCTGTGAGAGCAAGAGAGCTGTAACCTCTAACCCTTGA 659  
OY 541 GAATTTGGATCTGATTTCTTGAAGATACCGTTAATAGCAACTTATTCAGTGGGA 600  
DB 660 GAATTTGGATCTGATTTCTTGAAGATACCGTTAATAGCAACTTATTCAGTGGGA 719  
OY 601 GATCAAGATTTGTACAAATCACCCCTCAAGAAACCGGATGAATCACTTTGGATTC 660  
DB 720 GATCAAGATTTGTACAAATCACCCCTCAAGAAACCGGATGAATCACTTTGGATTC 779

OY 661 GCAAAAAAGGCTGCTGTGATTTTCTGAGAGGATGTAAACAATTAAGTCAATCATCA 720  
DB 780 GCAAAAAAGGCTGCTGTGATTTTCTGAGAGGATGTAAACAATTAAGTCAATCATCA 839  
OY 721 CCCAGTATATGATTTTGAACACACCTGAGAGGCTGAGTGCAGAGCATCCAGAAAG 780  
DB 840 CCCAGTATATGATTTTGAACACACCTGAGAGGCTGAGTGCAGAGCATCCAGAAAG 899  
OY 781 TATCAGGCTGATCTGTTTCAAACTTCCATGTGAGGCAATGTGGCAAAATATCACTGCC 840  
DB 900 TATCAGGCTGATCTGTTTCAAACTTCCATGTGAGGCAATGTGGCAAAATATCACTGCC 959  
OY 841 AGCTCATTAAGCATGAGAACAGAGCTTATTAATCACTAAGACAGATAATGATGATA 900  
DB 960 AGCTCATTAAGCATGAGAACAGAGCTTATTAATCACTAAGACAGATAATGATGATA 1019  
OY 901 AAGGCTGAATTTCTGTAATAAAGCAAAACGCTGGCTTACAGAGGAGCCAAATACAGA 960  
DB 1020 AAGGCTGAATTTCTGTAATAAAGCAAAACGCTGGCTTACAGAGGAGCCAAATACAGA 1079  
OY 961 TGGGCTGGAAGTAAAGAAACATGTAATGATAGGCGGAGCTCCACAGAGAAAAAGGTA 1020  
DB 1080 TGGGCTGGAAGTAAAGAAACATGTAATGATAGGCGGAGCTCCACAGAGAAAAAGGTA 1139  
OY 1021 GATCTGAATGCTGATCCCTGTGTGAGAGAAAAAAGTAATAGAGAAACTGCCATGC 1080  
DB 1140 GATCTGAATGCTGATCCCTGTGTGAGAGAAAAAAGTAATAGAGAAACTGCCATGC 1199  
OY 1081 TCAGAGATCCCTAGACATCTGAGAGATGTTCTTGATTAACACTTAATAGCAGATTAC 1140  
DB 1200 TCAGAGATCCCTAGACATCTGAGAGATGTTCTTGATTAACACTTAATAGCAGATTAC 1259  
OY 1141 AAAATTAATGAGTGTGTTTCCAGAGAGTGAATGATGATGTTGATGATGATGATGAT 1200  
DB 1260 AAAATTAATGAGTGTGTTTCCAGAGAGTGAATGATGATGATGATGATGATGATGAT 1319  
OY 1201 GGGGAGTCTGATCAATCAATGCGCAAAAGTGTGATGATGATGATGATGATGATGAT 1260  
DB 1320 GGGGAGTCTGATCAATCAATGCGCAAAAGTGTGATGATGATGATGATGATGATGAT 1379  
OY 1261 GAATATTTCTGTTCTTACAGAGAAATAGACTTACTGAGGAGTATCTCATAGGCTTTTA 1320  
DB 1380 GAATATTTCTGTTCTTACAGAGAAATAGACTTACTGAGGAGTATCTCATAGGCTTTTA 1439  
OY 1321 ATATGTAAGTGAAGAGTTCCTCAAGATCAATCAATCAATCAATCAATCAATCAAT 1380  
DB 1440 ATATGTAAGTGAAGAGTTCCTCAAGATCAATCAATCAATCAATCAATCAATCAAT 1499  
OY 1381 TTTGGGAAAACCTATCGGAAGAGGCAAGGCTCCCAACTTAAGCCATGTAAGTGA 1440  
DB 1500 TTTGGGAAAACCTATCGGAAGAGGCAAGGCTCCCAACTTAAGCCATGTAAGTGA 1559  
OY 1441 CTAAATTAAGAGCATTTGTTACTGAGCCACAGATTAATACAGAGGCTCCCTCACAAT 1500  
DB 1560 CTAAATTAAGAGCATTTGTTACTGAGCCACAGATTAATACAGAGGCTCCCTCACAAT 1619  
OY 1501 AAATTAAGGTTAAAGAGAGCCTATATCAAGGCTTATCCTGAGATTTTATCAAAA 1560  
DB 1620 AAATTAAGGTTAAAGAGAGCCTATATCAAGGCTTATCCTGAGATTTTATCAAAA 1679  
OY 1561 GCAGATTTGGCAGTTCAAAAGACTCTGAATATGATAATCAGGAACTAATCAACAGGAG 1620  
DB 1680 GCAGATTTGGCAGTTCAAAAGACTCTGAATATGATAATCAGGAACTAATCAACAGGAG 1739  
OY 1621 CAGATGCTCAAGTGAATGATATTAATTAATAGTGTATGAGATTAACAAAAGGTGAT 1680  
DB 1740 CAGATGCTCAAGTGAATGATATTAATTAATAGTGTATGAGATTAACAAAAGGTGAT 1799  
OY 1681 TCTATTAGAAATGGAAGAAATCCATACCAATGAAATCACTCGAANAAGATTCGCTTTC 1740  
DB 1800 TCTATTAGAAATGGAAGAAATCCATACCAATGAAATCACTCGAANAAGATTCGCTTTC 1859  
OY 1741 AAAAGCAAGCTGAACTATAGCAGCATATAAGCAATATGAACTGAAATTAATATC 1800



```
|||||
Db 1860 AAAACAAAGCTACCTATATACGACGATATAGCATATGGAATCGAATTAATATC 1919
OY 1801 CACATTTCAAAGACACCTATAAAGAAATAGGCTGAGGAGAGAGCTTCTACAGGCATAT 1860
Db 1920 CACATTTCAAAGACACCTATAAAGAAATAGGCTGAGGAGAGAGCTTCTACAGGCATAT 1979
OY 1861 CATGCCCTTGAACTAGTAGTACGTAGTAAGAAATCTAAGCCACCTAATTTAGTGAATTC 1920
Db 1980 CATGCCCTTGAACTAGTAGTACGTAGTAAGAAATCTAAGCCACCTAATTTAGTGAATTC 2039
OY 1921 ATTGAAGTTGTTCTAGCACTGAGATGAAGATTAAGAAAAAGTCAACCAATGCCAGTC 1980
Db 2040 ATTGAAGTTGTTCTAGCACTGAGATGAAGATTAAGAAAAAGTCAACCAATGCCAGTC 2099
OY 1981 AGGCACAGCAAAACCTACAACTCATGGAAGTAAAGAACCTGCACTGAGCCAAAG 2040
Db 2100 AGGCACAGCAAAACCTACAACTCATGGAAGTAAAGAACCTGCACTGAGCCAAAG 2159
OY 2041 AGTACAGCCAAATGACAGACAGTAAAAAGACATGACAGTACTTTCCAGAGCTG 2100
Db 2160 AGTACAGCCAAATGACAGACAGTAAAAAGACATGACAGTACTTTCCAGAGCTG 2219
OY 2101 AAGTTACAAATGACACCTGGTCTTTTACTAAGTTCATTAAGTTCACAGTGAATTAAG 2160
Db 2220 AAGTTACAAATGACACCTGGTCTTTTACTAAGTTCATTAAGTTCACAGTGAATTAAG 2279
OY 2161 TTTGTCAATCTAGCTTCCAAAGAGAGAAAAAGAGAACTAGAAACAGTTAAAGT 2220
Db 2280 TTTGTCAATCTAGCTTCCAAAGAGAGAAAAAGAGAACTAGAAACAGTTAAAGT 2339
OY 2221 TCTAATATGCTGAGAGACCCCAAGATCTCAGTTAAGTGAAGAGGGTTTGCACACT 2280
Db 2340 TCTAATATGCTGAGAGACCCCAAGATCTCAGTTAAGTGAAGAGGGTTTGCACACT 2299
OY 2281 GAAAGATCTGTGAGAGTACGAGTATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 2340
Db 2400 GAAAGATCTGTGAGAGTACGAGTATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2459
OY 2341 GAAAGATCTGTGAGAGTACGAGTATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400
Db 2460 GAAAGATCTGTGAGAGTACGAGTATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2519
OY 2401 TGTGTGAGTCACTGTGACAGCATTTTGAAGCCCAAGGACTAATTCATGCTGCTGCTGCT 2460
Db 2520 TGTGTGAGTCACTGTGACAGCATTTTGAAGCCCAAGGACTAATTCATGCTGCTGCTGCTGCT 2579
OY 2461 GATAATAGAAATGACACAGAGGCTTAAATGATCCATTTGGACATGAGTTAACCCAGT 2520
Db 2580 GATAATAGAAATGACACAGAGGCTTAAATGATCCATTTGGACATGAGTTAACCCAGT 2639
OY 2521 CGGGAACAGCATAGAAATGAGAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2580
Db 2640 CGGGAACAGCATAGAAATGAGAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2699
OY 2581 TTTCAAGTTTCAAGCGCAGTCAATTTGCTGCTTTTCAAAATCAGAGAAATGACAGAGAG 2640
Db 2700 TTTCAAGTTTCAAGCGCAGTCAATTTGCTGCTTTTCAAAATCAGAGAAATGACAGAGAG 2759
OY 2641 GAATGTGCAACATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700
Db 2760 GAATGTGCAACATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2819
OY 2701 TTTGAATGTGAACAAAGAGAAATCAAGAAAGAAATGATGATTAATCAAGCTGTA 2760
Db 2820 TTTGAATGTGAACAAAGAGAAATCAAGAAAGAAATGATGATTAATCAAGCTGTA 2879
OY 2761 CAGACAGTTAATATGACAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820
Db 2880 CAGACAGTTAATATGACAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2939
OY 2821 AATGCCAATGTAGTATCAAGAGAGGCTCTAGGTTTGTCTATCATCTCAGTTCCAGAGC 2880
|||||
|||||
Db 2940 AATGCCAATGTAGTATCAAGAGAGGCTCTAGGTTTGTCTATCATCTCAGTTCCAGAGC 2999
OY 2881 AAGGAACCTGAGCTCTTACTCCAAATTAACATGAGCTTTTCAAAACCCATATGTA 2940
Db 3000 AAGGAACCTGAGCTCTTACTCCAAATTAACATGAGCTTTTCAAAACCCATATGTA 3059
OY 2941 CCACACCTTTTCCATCAAGTCAATTTGTTAAACCAAAATGTAAGAAAAATCTGCTAAG 3000
Db 3060 CCACACCTTTTCCATCAAGTCAATTTGTTAAACCAAAATGTAAGAAAAATCTGCTAAG 3119
OY 3001 GAAACCTTTGAGAACATTCATGTCACCTGGAAGAGAAATGGAATGAGAACATTTCCA 3060
Db 3120 GAAACCTTTGAGAACATTCATGTCACCTGGAAGAGAAATGGAATGAGAACATTTCCA 3179
OY 3061 AGTACAGTGAAGCAATTTAGCCGTATTAATACATTTAGGAAAAATGTTTAAAGAACCCAGC 3120
Db 3180 AGTACAGTGAAGCAATTTAGCCGTATTAATACATTTAGGAAAAATGTTTAAAGAACCCAGC 3239
OY 3121 TCAAGCAATTAATGAAGTGAAGTTCACAGTAAATGAAGTGGCTCCAGTATTAATGA 3180
Db 3240 TCAAGCAATTAATGAAGTGAAGTTCACAGTAAATGAAGTGGCTCCAGTATTAATGA 3299
OY 3181 ATAGGTTCCAGTATGATAAAAACATTCAGACGAACTAGTGAAGACAGAGGCCAAATTC 3240
Db 3300 ATAGGTTCCAGTATGATAAAAACATTCAGACGAACTAGTGAAGACAGAGGCCAAATTC 3359
OY 3241 AATGCTATGCTAGATTAAGGGGTTTGGCACTGAGGCTGATTAACAAAGTCTTCCGGA 3300
Db 3360 AATGCTATGCTAGATTAAGGGGTTTGGCACTGAGGCTGATTAACAAAGTCTTCCGGA 3419
OY 3301 AGTAATTTGAACATCTGAAATTAAGAAAGCAAGATATGAAGAGTATGATGATGATGAT 3360
Db 3420 AGTAATTTGAACATCTGAAATTAAGAAAGCAAGATATGAAGAGTATGATGATGATGAT 3479
OY 3361 AATACAGATTTTCTGCAATATCTGATTTCAATTAAGACAGCCTATGGAAGTACT 3420
Db 3480 AATACAGATTTTCTGCAATATCTGATTTCAATTAAGACAGCCTATGGAAGTACT 3539
OY 3421 CATGATCTCAGGTTTGTCTGAGACACCTGATGCTGTTAAGATGATGATGATGATGAT 3480
Db 3540 CATGATCTCAGGTTTGTCTGAGACACCTGATGCTGTTAAGATGATGATGATGATGAT 3599
OY 3481 GAAAGATCTAGTTTGTCTGAAATGACATTAAGAAAGTCTGCTGTTTTCAGAAAAAGC 3540
Db 3600 GAAAGATCTAGTTTGTCTGAAATGACATTAAGAAAGTCTGCTGTTTTCAGAAAAAGC 3659
OY 3541 GTCCAGAAAGGAGACCTTAAGCAGAGTCTTACCCCTTACCCATACATTTGGCTCAG 3600
Db 3660 GTCCAGAAAGGAGACCTTAAGCAGAGTCTTACCCCTTACCCATACATTTGGCTCAG 3719
OY 3601 GGTACCGAAGAGGGGCCAAGAAATTAAGTCTCAGAGAACTTAATCTAGTGAAGAT 3660
Db 3720 GGTACCGAAGAGGGGCCAAGAAATTAAGTCTCAGAGAACTTAATCTAGTGAAGAT 3779
OY 3661 GAAAGCTTCCCTGCTTCCAACTGTTATTTGGTAAAGTAAACAATATACCTTCAG 3720
Db 3780 GAAAGCTTCCCTGCTTCCAACTGTTATTTGGTAAAGTAAACAATATACCTTCAG 3839
OY 3721 TCTACTAGGCAATGACAGGCTTCTACGAGTGTCTGTAAAGACACAGAGAGAAATTA 3780
Db 3840 TCTACTAGGCAATGACAGGCTTCTACGAGTGTCTGTAAAGACACAGAGAGAAATTA 3899
OY 3781 TTATCATTTGAAGATAGCTTAAATAGTCAAGTCAACCGATTAATTTGGCAAGGCACT 3840
Db 3900 TTATCATTTGAAGATAGCTTAAATAGTCAAGTCAACCGATTAATTTGGCAAGGCACT 3959
OY 3841 CAGGAACATCACTAGTGAAGAAACAAATGTTCTGCTAGCTGTTTCTTCCAGATGC 3900
Db 3960 CAGGAACATCACTAGTGAAGAAACAAATGTTCTGCTAGCTGTTTCTTCCAGATGC 4019
OY 3901 AGTGAATTTGAAGACTTCACTGCAAAATCAAAACCCAGATCCCTTCTGATTTGATGCT 3960
Db 4020 AGTGAATTTGAAGACTTCACTGCAAAATCAAAACCCAGATCCCTTCTGATTTGATGCT 4079
```

QY 3961 TCACAAATATGAGGATCAGTCTGAAAGCCAGGAGTGTCTGAGTGAACAAGAAATG 4020  
D 4080 TCACAAATATGAGGATCAGTCTGAAAGCCAGGAGTGTCTGAGTGAACAAGAAATG 4139  
QY 4021 GTTTCATATATGAGAAAGAGAGACGGCTTGGAAAGAAATATATCAAGAGCAAGC 4080  
D 4140 GTTTCATATATGAGAAAGAGAGACGGCTTGGAAAGAAATATATCAAGAGCAAGC 4199  
QY 4081 ATGATTCATCAACTAGTGTGAAGCAGCATCTGGGTGTGAGAGTGAACAACAGCTCTGAA 4140  
D 4200 ATGATTCATCAACTAGTGTGAAGCAGCATCTGGGTGTGAGAGTGAACAACAGCTCTGAA 4259  
QY 4141 GACTGTCTAGGGCTATCTCTCTCAGAGTGAATTTTAAACACTCAGCAGAGGATACCATG 4200  
D 4260 GACTGTCTAGGGCTATCTCTCTCAGAGTGAATTTTAAACACTCAGCAGAGGATACCATG 4319  
QY 4201 CAACATTAACCTGAATTAAGCTCCAGCAAGAAATGGCTCACTAGAGCTGTGTGAACAG 4280  
D 4320 CAACATTAACCTGAATTAAGCTCCAGCAAGAAATGGCTCACTAGAGCTGTGTGAACAG 4379  
QY 4261 CATGGAGCCAGCCTTCTTAACAGTACCTTCCATCATTAAGTGAATCTTCTGCCCTTGAG 4320  
D 4380 CATGGAGCCAGCCTTCTTAACAGTACCTTCCATCATTAAGTGAATCTTCTGCCCTTGAG 4439  
QY 4321 GACCTGCAAAATCCAGAACAGCATCAGAAAAAGCAGTATTAACTTCACAGAAAAAGT 4380  
D 4440 GACCTGCAAAATCCAGAACAGCATCAGAAAAAGCAGTATTAACTTCACAGAAAAAGT 4499  
QY 4381 AGTGAATACCTCTATTAAGCAGATCCAGAAAGCCTTCTCTGTCAGCAAGTTTGAAGTCT 4440  
D 4500 AGTGAATACCTCTATTAAGCAGATCCAGAAAGCCTTCTCTGTCAGCAAGTTTGAAGTCT 4559  
QY 4441 GCAGATAGTTCTACCACTAAATAAAGAACAGAGAGTGAAGAGTATCCCTTCTTAA 4500  
D 4560 GCAGATAGTTCTACCACTAAATAAAGAACAGAGAGTGAAGAGTATCCCTTCTTAA 4619  
QY 4501 TGCCCATCATTAAGTGAATGAGTGTGATCAGACAGTTGCTGTGGAGTCTTCAGAAATGA 4560  
D 4620 TGCCCATCATTAAGTGAATGAGTGTGATCAGACAGTTGCTGTGGAGTCTTCAGAAATGA 4679  
QY 4561 AACACCCATCTCAAGAGAGCCTCATTAAGTGTGTATATGTGAGAGGACAAAGCTGGAA 4620  
D 4680 AACACCCATCTCAAGAGAGCCTCATTAAGTGTGTATATGTGAGAGGACAAAGCTGGAA 4739  
QY 4621 GAGTCTGGGCAACAGATTTGAGCGAAACATCTTACTTGGCAAGGCAAGATCTAGAGGA 4680  
D 4740 GAGTCTGGGCAACAGATTTGAGCGAAACATCTTACTTGGCAAGGCAAGATCTAGAGGA 4799  
QY 4681 ACCCCTTAACCTGAATCTGAATCAGCCTCTCTGTGATGAGACCCCTGAATCTGATCTTCT 4740  
D 4800 ACCCCTTAACCTGAATCTGAATCAGCCTCTCTGTGATGAGACCCCTGAATCTGATCTTCT 4859  
QY 4741 GAGAGACGAGGCCCCAGAGTACGTCGTTGGCAACATTAACCTTCACTCTGATG 4800  
D 4860 GAGAGACGAGGCCCCAGAGTACGTCGTTGGCAACATTAACCTTCACTCTGATG 4919  
QY 4801 AAAGTTCCCAATTAAGTGTGAAGATCTGCCAGAGTCCAGCTGCTGCTACTACT 4860  
D 4920 AAAGTTCCCAATTAAGTGTGAAGATCTGCCAGAGTCCAGCTGCTGCTACTACT 4979  
QY 4861 GATACCTGCTGGATTAATGCAATGAGAAAGTGTGAGAGGAGAAAGCAGAAATTGACA 4920  
D 4980 GATACCTGCTGGATTAATGCAATGAGAAAGTGTGAGAGGAGAAAGCAGAAATTGACA 5039  
QY 4921 GCTTCAACAGAAAGGCAACAAAGAAATGTCCATGCTGTGCTGCCCTGAGCCCGAGAA 4980  
D 5040 GCTTCAACAGAAAGGCAACAAAGAAATGTCCATGCTGTGCTGCCCTGAGCCCGAGAA 5099  
QY 4981 GAATTTATGCTGCTGCTCAAGTTTGGCAGAAACACACATCACTTAATTAATCTAAT 5040  
D 5100 GAATTTATGCTGCTGCTCAAGTTTGGCAGAAACACACATCACTTAATTAATCTAAT 5159

QY 5041 ACTGAGAGACTACTATGTTGTTATGAAACAGATCTGAGTTGTGTGTAACGAGACA 5100  
D 5160 ACTGAGAGACTACTATGTTGTTATGAAACAGATCTGAGTTGTGTGTAACGAGACA 5219  
QY 5101 CTGAAATATTTCTAGGAATTTGGGGAGAAATGGGTACTGTGCTTCTGGGTGACC 5160  
D 5220 CTGAAATATTTCTAGGAATTTGGGGAGAAATGGGTACTGTGCTTCTGGGTGACC 5279  
QY 5161 CAGTCTATTAAGAAAGAAATGCTGAATGAGCATATTTTGAAGTCAGAGAGATGTG 5220  
D 5280 CAGTCTATTAAGAAAGAAATGCTGAATGAGCATATTTTGAAGTCAGAGAGATGTG 5339  
QY 5221 GTCATGGAAGAAACCAAGAGTCCAAAGCGAGCAAGAAATCCAGACAGAAAGATC 5280  
D 5340 GTCATGGAAGAAACCAAGAGTCCAAAGCGAGCAAGAAATCCAGACAGAAAGATC 5399  
QY 5281 TTCCAGGGGCTAGAAATCTGTGTATGGGCGCTTCACCAACATGGCCACAGATCACTG 5340  
D 5400 TTCCAGGGGCTAGAAATCTGTGTATGGGCGCTTCACCAACATGGCCACAGATCACTG 5459  
QY 5341 GAATGATGTTAGAGCTGTGTGTGCTTCTGTGTGAGAGAGCTTTCATCATTCACCTT 5400  
D 5460 GAATGATGTTAGAGCTGTGTGTGCTTCTGTGTGAGAGAGCTTTCATCATTCACCTT 5519  
QY 5401 GGCAAGGTGTCCACCAATTTGTGTTGTGCAGCCAGATGCTTGAGACAGACAAATGTC 5460  
D 5520 GGCAAGGTGTCCACCAATTTGTGTTGTGTGCAGCCAGATGCTTGAGACAGACAAATGTC 5579  
QY 5461 TTCCATGCAATTTGGGCGAGATGTGTGAGGACCTGTGTGACCCGAGAGTGGGTGTGAC 5520  
D 5580 TTCCATGCAATTTGGGCGAGATGTGTGAGGACCTGTGTGACCCGAGAGTGGGTGTGAC 5639  
QY 5521 AGTGTACACTCTACAGTGCAGAGAGCTGGAGACCTGATATACCCAGATCCCCAC 5580  
D 5640 AGTGTACACTCTACAGTGCAGAGAGCTGGAGACCTGATATACCCAGATCCCCAC 5699  
QY 5581 AGCCACTAC 5589  
D 5700 AGCCACTAC 5708

RESULT 5  
US-10-022-819-1  
Sequence 1, Application US/10022819  
Publication No. US20030027166A1  
GENERAL INFORMATION:  
APPLICANT: ALLEN, Antonette C. P.  
LAWRENCE, Tammy  
ANGELLY, Tracy S.  
RABIN, Mark B.  
TITLE OF INVENTION: CODING SEQUENCE HAPLOTYPE OF THE HUMAN  
BRCA1 GENE  
NUMBER OF SEQUENCES: 67  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morgan Lewis & Bockius LLP  
STREET: 1111 Pennsylvania Avenue  
CITY: Washington DC  
STATE: District of Columbia  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/022,819  
FILING DATE: 22-Apr-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/074,452  
FILING DATE: 1998-05-06

ATTORNEY/AGENT INFORMATION:  
NAME: <Unknown>  
REGISTRATION NUMBER: <Unknown>  
REFERENCE/DOCKET NUMBER: 044921-5049-01-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-739-3001  
TELEFAX: 202-739-3000  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5711 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: HOMO SAPIENS  
STRAIN: BRCA1  
HAPOTYPE: OMT4  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: 17  
MAP POSITION: 17q21  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-10-022-819-1  
  
Query Match 99.9%; Score 5582.6; DB 9; Length 5711.  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 5585; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
  
QY 1 ATGAGTTTATCTGCTCTGCGCTTGAAGAAGTACAAATGCTTAATGCTATGACAGAA 60  
DB 120 ATGAGTTTATCTGCTCTGCGCTTGAAGAAGTACAAATGCTTAATGCTATGACAGAA 179  
QY 61 ATCTTAGAGTGTCCCATCTGTCTGGAGTTGATCAAGSACCTGTCTCCAAAGTGTGAC 120  
DB 180 ATCTTAGAGTGTCCCATCTGTCTGGAGTTGATCAAGSACCTGTCTCCAAAGTGTGAC 239  
QY 121 CACATATTTTGCAAATTTTGCATGCTGAAACCTCTCAACAGAGAAAGGCGCTTCACAG 180  
DB 240 CACATATTTTGCAAATTTTGCATGCTGAAACCTCTCAACAGAGAAAGGCGCTTCACAG 299  
QY 181 TGCTCTTATGTAAGAATATATACCAAAAGAGCGCTACAGAAAGTACAGATTTAGT 240  
DB 300 TGCTCTTATGTAAGAATATATACCAAAAGAGCGCTACAGAAAGTACAGATTTAGT 359  
QY 241 CAACTTGTGAAGAGCTATTTGAATCATTTTGTCTTTTACGCTTGACACAGGTTTGGAG 300  
DB 360 CAACTTGTGAAGAGCTATTTGAATCATTTTGTCTTTTACGCTTGACACAGGTTTGGAG 419  
QY 301 TATGCAAAAGAGCTATATTTTGCAAAAGAAAGAAATTAACCTCTCTGAAACATCTAAAGAT 360  
DB 420 TATGCAAAAGAGCTATATTTTGCAAAAGAAAGAAATTAACCTCTCTGAAACATCTAAAGAT 479  
QY 361 GAAGTTTCTATCATCCAAAGATAGGCTACAGAAACCGTGCCAAAAGACTTCTACAGAGT 420  
DB 480 GAAGTTTCTATCATCCAAAGATAGGCTACAGAAACCGTGCCAAAAGACTTCTACAGAGT 539  
QY 421 GAACCCGAAAATCTCTTCTTGACAGAAACAGTCTCAGTCCCAACTCTTAACCTTTGA 480  
DB 540 GAACCCGAAAATCTCTTCTTGACAGAAACAGTCTCAGTCCCAACTCTTAACCTTTGA 599  
QY 481 ACTGTGAGAACTCTGAGCAAAAGACGCGATACCAACTCAAAAGAGCTGTGCTACAT 540  
DB 600 ACTGTGAGAACTCTGAGCAAAAGACGCGATACCAACTCAAAAGAGCTGTGCTACAT 659  
QY 541 GAATGGAGTCTGATTTCTTGAAGATACCGTTAATAGCAACTTAATGACAGTGTGGGA 600  
DB 660 GAATGGAGTCTGATTTCTTGAAGATACCGTTAATAGCAACTTAATGACAGTGTGGGA 719  
QY 601 GATCAAGAAATTTGTACAAATCACCCCTCAAGGAACAGGAGTGAATCAGTTTGATTC 660  
|||||

DB 720 GATCAAGAAATTTGTACAAATCACCCCTCAAGGAACAGGAGTGAATCAGTTTGATTC 779  
QY 661 GCAAAAAGAGCTCTGTTGTAATTTTCTGAGAGGATGTAAACAATATGTAACATGATCA 720  
DB 780 GCAAAAAGAGCTCTGTTGTAATTTTCTGAGAGGATGTAAACAATATGTAACATGATCA 839  
QY 721 CCCAGTAATTAATGATTTGAACACCACTGAGAGCGTGCAGCTGAGAGGATCCAGAAA 780  
DB 840 CCCAGTAATTAATGATTTGAACACCACTGAGAGCGTGCAGCTGAGAGGATCCAGAAA 899  
QY 781 TATCAGGAGTGTCTGTTTCAAACTTGATGAGAGCCATGTGGCAACAATATCATGCC 840  
DB 900 TATCAGGAGTGTCTGTTTCAAACTTGATGAGAGCCATGTGGCAACAATATCATGCC 959  
QY 841 AGCTCATTTACAGATGAGAAACAGCACTTAATTAATCTAATAAGACAGATGATGTAGAA 900  
DB 960 AGCTCATTTACAGATGAGAAACAGCACTTAATTAATCTAATAAGACAGATGATGTAGAA 1019  
QY 901 AAGGCTGAATTCGTATATAAAGCAAAACAGCCTGGCTTAGCAAGAGCCCAACATTAACA 960  
DB 1020 AAGGCTGAATTCGTATATAAAGCAAAACAGCCTGGCTTAGCAAGAGCCCAACATTAACA 1079  
QY 961 TGGGCTGGAAGTGAAGAAACATGTATGATGAGCGGACCTCCAGCAGCAAGAAAAAGTA 1020  
DB 1080 TGGGCTGGAAGTGAAGAAACATGTATGATGAGCGGACCTCCAGCAGCAAGAAAAAGTA 1139  
QY 1021 GATCTGAATGCTGATCCCTGTTGAGAGAAAAAGATGGAATTAACAGAACTGCCATGC 1080  
DB 1140 GATCTGAATGCTGATCCCTGTTGAGAGAAAAAGATGGAATTAACAGAACTGCCATGC 1199  
QY 1081 TCAGAGAAATCTGAGATACTGAAAGATGTCTTGATTAACATAAATAGCAGATTGAG 1140  
DB 1200 TCAGAGAAATCTGAGATACTGAAAGATGTCTTGATTAACATAAATAGCAGATTGAG 1259  
QY 1141 AAGGTTAATGAGTGTGTTTCCAGAAAGTGAAGTGTGTTGATGATGATGATGATGAT 1200  
DB 1260 AAGGTTAATGAGTGTGTTTCCAGAAAGTGAAGTGTGTTGATGATGATGATGATGAT 1319  
QY 1201 GGGAGCTGTAATCAATCAATGCAAGTGAAGTGTGTTGATGATGATGATGATGATGAT 1260  
DB 1320 GGGAGCTGTAATCAATCAATGCAAGTGAAGTGTGTTGATGATGATGATGATGATGAT 1379  
QY 1261 GAATATTTCTGTTCTTCAAGAAATATAGACTTACGCGCAGATTCCTCATGAGCTTTA 1320  
DB 1380 GAATATTTCTGTTCTTCAAGAAATATAGACTTACGCGCAGATTCCTCATGAGCTTTA 1439  
QY 1321 ATATGTAAGTGAAGAGCTTCACTCCAAATGATGAGATGATGATGATGATGATGAT 1380  
DB 1440 ATATGTAAGTGAAGAGCTTCACTCCAAATGATGAGATGATGATGATGATGATGAT 1499  
QY 1381 TTTGGGAAAACCTATGGAAGAGCAAGCTCCCACTTAAGCCATGTAAGTGAAT 1440  
DB 1500 TTTGGGAAAACCTATGGAAGAGCAAGCTCCCACTTAAGCCATGTAAGTGAAT 1559  
QY 1441 CTAAATTAAGAGAGCTTTGTTACTGAGCCAGATTAATCAAGAGCTGCCCTACAAT 1500  
DB 1560 CTAAATTAAGAGAGCTTTGTTACTGAGCCAGATTAATCAAGAGCTGCCCTACAAT 1619  
QY 1501 AAATTAAGAGTGAAGAGAGCTTCAATCAGGCGCTTCAATCAGAGATTTTATCAAGAA 1560  
DB 1620 AAATTAAGAGTGAAGAGAGCTTCAATCAGGCGCTTCAATCAGAGATTTTATCAAGAA 1679  
QY 1561 GCAGATTTGGCAGTTCAAAAGACTCCTGAAATGATTAATCAGGAACTTAACCAAGGAG 1620  
DB 1680 GCAGATTTGGCAGTTCAAAAGACTCCTGAAATGATTAATCAGGAACTTAACCAAGGAG 1739  
QY 1621 CAGAAATGTCAGTGAATGATTAATTAATGATGATGATGATGATGATGATGATGAT 1680  
DB 1740 CAGAAATGTCAGTGAATGATTAATTAATGATGATGATGATGATGATGATGATGAT 1739  
QY 1681 TCTATTCAAGATGAGAAAAATCTTAACCAATTAATGATGATGATGATGATGATGAT 1740  
DB 1800 TCTATTCAAGATGAGAAAAATCTTAACCAATTAATGATGATGATGATGATGATGAT 1859  
|||||

1741 AAAACGAAGCTGAACCTATTAAGCAGCATATAAGCAATATGGAATCGAATTAATATC 1800  
1860 AAAACGAAGCTGAACCTATTAAGCAGCATATAAGCAATATGGAATCGAATTAATATC 1919  
1801 CACAAATTCAAAAGCAGCTTAAGAAATAGCCTGAGGAGAAAGTCTTCTACCGAGCATATT 1860  
1920 CACAAATTCAAAAGCAGCTTAAGAAATAGCCTGAGGAGAAAGTCTTCTACCGAGCATATT 1979  
1861 CATGCGCTTGAACAGTAGTAGTAGTAAGTAATTAAGCCCACTAATTTGACTGAATTGCA 1920  
1980 CATGCGCTTGAACAGTAGTAGTAGTAAGTAATTAAGCCCACTAATTTGACTGAATTGCA 2039  
1921 ATTGATATGTTGTTTACAGTAGAGAGATTAAGAAAAAAAGTCAACCAATATCCAGTC 1980  
2040 ATTGATATGTTGTTTACAGTAGAGAGATTAAGAAAAAAAGTCAACCAATATCCAGTC 2039  
1981 AGGCACAGCAGAAACCTACAACTCATGGAAGGTAAAGAACCTGCACTGGAGCCAGAAG 2040  
2100 AGGCACAGCAGAAACCTACAACTCATGGAAGGTAAAGAACCTGCACTGGAGCCAGAAG 2159  
2041 AGTACAGCGCAATTAAGACAGATTAAGAAAGATACATGATGATCTTCTCCAGACTG 2100  
2160 AGTACAGCGCAATTAAGACAGATTAAGAAAGATACATGATGATCTTCTCCAGACTG 2219  
2101 AAGTTACAATATGACCTGCTGTTCTTACTAAGTGTCAATATACAGTGAACCTTAAGAA 2160  
2220 AAGTTACAATATGACCTGCTGTTCTTACTAAGTGTCAATATACAGTGAACCTTAAGAA 2279  
2161 TTTGTCAATCTTACGCTTCCAAAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGTAAAGT 2220  
2280 TTTGTCAATCTTACGCTTCCAAAGAGAAAGAAAGAAAGAAAGAAAGAAAGTAAAGT 2339  
2221 TCTAATATGCTGAAGCCCCAAAGATCTCATGTTAAGTGAAGAAAGGTTTGGCAACT 2280  
2340 TCTAATATGCTGAAGCCCCAAAGATCTCATGTTAAGTGAAGAAAGGTTTGGCAACT 2399  
2281 GAAAGATCTGTAGAGATAGCAGATTTTCAATGCTGACTGATGATTAAGGACTCAG 2340  
2400 GAAAGATCTGTAGAGATAGCAGATTTTCAATGCTGACTGATGATTAAGGACTCAG 2459  
2341 GAAAGTATCTGCTTACTGAGATTAGCAGTCTAGGAGAAAGCAAAACAGAACCAATATA 2400  
2460 GAAAGTATCTGCTTACTGAGATTAGCAGTCTAGGAGAAAGCAAAACAGAACCAATATA 2519  
2401 TGTGTGAGTGTGAGATTTGAACCCCAAGGACTAATCTATGTTGTTCCAAA 2460  
2520 TGTGTGAGTGTGAGATTTGAACCCCAAGGACTAATCTATGTTGTTCCAAA 2579  
2461 GATTAATGAATGACACAGAGGCTTTAATATCATTGGACATGAAAGTTAACACACAGT 2520  
2580 GATTAATGAATGACACAGAGGCTTTAATATCATTGGACATGAAAGTTAACACACAGT 2639  
2521 CGGGAACAACAGATAGAAATGAGAAAGTGAACCTGATCTCAGTATTTGACGAATACA 2580  
2640 CGGGAACAACAGATAGAAATGAGAAAGTGAACCTGATCTCAGTATTTGACGAATACA 2699  
2581 TTCAAGGTTCAAGGCCGAGTATTTGCTGTTTCAATCCAGAAATGAGAAAG 2640  
2700 TTCAAGGTTCAAGGCCGAGTATTTGCTGTTTCAATCCAGAAATGAGAAAG 2759  
2641 GAATGTGCAATTTCTGCCCACTGAGGCTTTAAGAAACAAGATCCAAAAGTACT 2700  
2760 GAATGTGCAATTTCTGCCCACTGAGGCTTTAAGAAACAAGATCCAAAAGTACT 2819  
2701 TTTGAATGTGACAAAAGAGAAATCAAGAAAGATGATCTTAATATCAAGCCCTGA 2760  
2820 TTTGAATGTGACAAAAGAGAAATCAAGAAAGATGATCTTAATATCAAGCCCTGA 2879  
2761 CAGACGTTATATACACTGAGGCTTCTGTTGCTGTCAGAAAGATAGCCAGTTGAT 2820  
2880 CAGACGTTATATACACTGAGGCTTCTGTTGCTGTCAGAAAGATAGCCAGTTGAT 2939

2821 AATGCCAATGTAGTATCAAGAGAGGCTAGGTTTGTCTATCATCTCACTTCAGAGGC 2880  
2940 AATGCCAATGTAGTATCAAGAGAGGCTAGGTTTGTCTATCATCTCACTTCAGAGGC 2999  
2881 AACGAACCTGACATCTTACTCCAAATTAACATGAGCTTTTACAAAACCATATGATTA 2940  
3000 AACGAACCTGACATCTTACTCCAAATTAACATGAGCTTTTACAAAACCATATGATTA 3059  
2941 CCACCACTTTTCCCATCAAGTCACTTTGTATAAATCTAAATGTAGAAAAATCTCTAG 3000  
3060 CCACCACTTTTCCCATCAAGTCACTTTGTATAAATCTAAATGTAGAAAAATCTCTAG 3119  
3001 GAAACCTTTGAGAAACATTAATATGACCTGAAGAGAAATGGAATGAGAACATTCGA 3060  
3120 GAAACCTTTGAGAAACATTAATATGACCTGAAGAGAAATGGAATGAGAACATTCGA 3179  
3061 AGTACAGTGAACCAATTAAGCCGTAATTAACATTAAGAAAAATGTTTAAAGAACCCAGC 3120  
3180 AGTACAGTGAACCAATTAAGCCGTAATTAACATTAAGAAAAATGTTTAAAGAACCCAGC 3239  
3121 TCAAGCAATATTAATGAAGTAGTTCAGTACTAATGAAGTGGCTCCAGTATTAATGA 3180  
3240 TCAAGCAATATTAATGAAGTAGTTCAGTACTAATGAAGTGGCTCCAGTATTAATGA 3299  
3181 ATAGGTTCCAGTGAAGAAACATTCAGAGCAAGAACTAGTGAAGAACAGAGGCCAAATTTG 3240  
3300 ATAGGTTCCAGTGAAGAAACATTCAGAGCAAGAACTAGTGAAGAACAGAGGCCAAATTTG 3359  
3241 AATGCTATGCTTATGATTAGGGGTTTCCAACTGAGGCTATATAAACAAGTCTCTGGA 3300  
3360 AATGCTATGCTTATGATTAGGGGTTTCCAACTGAGGCTATATAAACAAGTCTCTGGA 3419  
3301 AGTATATGTAAGCACTCTGAAATTAAGAAACAAATATGAAGAGTAGTTCAACCTGT 3360  
3420 AGTATATGTAAGCACTCTGAAATTAAGAAACAAATATGAAGAGTAGTTCAACCTGT 3479  
3361 AATACAGATTTCTCTCATATCTGATTTCAATTAAGAACAGCTATGGAAGTAGT 3420  
3480 AATACAGATTTCTCTCATATCTGATTTCAATTAAGAACAGCTATGGAAGTAGT 3559  
3421 CATGATCTCAGGTTTCTTGAGACACCTGATGACCTGTTAGATGATGGAATTAAG 3480  
3540 CATGATCTCAGGTTTCTTGAGACACCTGATGACCTGTTAGATGATGGAATTAAG 3599  
3481 GAAGATCTGATGTTTGTGAAATAGCATTAAGAAAGTCTGCTGTTTATGCAAAAGC 3540  
3600 GAAGATCTGATGTTTGTGAAATAGCATTAAGAAAGTCTGCTGTTTATGCAAAAGC 3659  
3541 GTCCAGAAAGAGAGCTTTACAGAGTCTAGGCCCTTCAACCCATACATTTGGCTCAG 3600  
3660 GTCCAGAGAGAGAGCTTTACAGAGTCTAGGCCCTTCAACCCATACATTTGGCTCAG 3719  
3601 GGTTCACGAAGAGGAGGCAAGAAATTAAGAGTCTCAGAGAGAACTTATCTAGTAGAT 3660  
3720 GGTTCACGAAGAGGAGGCAAGAAATTAAGAGTCTCAGAGAGAACTTATCTAGTAGAT 3779  
3780 GAAGAGCTTCTCTGCTTCCAACTGTTATTTGGTAAAGTAAACATTAATACCTTCTCAG 3839  
3721 TCTACTAGCATATGACACGTTGCTACCGAGTGTCTCTAAGAAACAGAGAGAAATTTA 3780  
3840 TCTACTAGCATATGACACGTTGCTACCGAGTGTCTCTAAGAAACAGAGAGAAATTTA 3899  
3781 TTATCATTTGAAGATATGCTTAATAGCTGCAATACAGGATTAATTTGGCAAGGATCT 3840  
3900 TTATCATTTGAAGATATGCTTAATAGCTGCAATACAGGATTAATTTGGCAAGGATCT 3959  
3841 CAGGAACATACCTTATGAGAGAAACAAATGTTCTGCTAGCTGTTTCTTCAACAGTGC 3900  
3960 CAGGAACATACCTTATGAGAGAAACAAATGTTCTGCTAGCTGTTTCTTCAACAGTGC 4019  
3901 AGTGAATTTGAAGACTTGAATCAAAATACAAACACCCAGATCTTCTGATTTGATCT 3960



PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/966,436  
FILING DATE: 07-NO: US20020183268A1-97  
APPLICATION NUMBER: US 08/598,591  
FILING DATE: 12-Feb-96  
ATTORNEY/AGENT INFORMATION:  
NAME: Michael S. Tuscan  
REGISTRATION NUMBER: 43,210  
REFERENCE/DOCKET NUMBER: 44921-5055-02-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-739-3000  
TELEFAX: 202-739-3001  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5711 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: No. US20020183268A1 Relevant  
TOPOLOGY: Linear  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
STRAIN: BRCAL  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: 17  
MAP POSITION: 17q21  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-734-672-1

Query Match 99.98; Score 5581; DB 9; Length 5711;  
Best Local Similarity 99.98; Pred. No. 0;  
Matches 5584; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

1 ATGATTTATCTGCTCTCTGCGTTGAGAGATCAAAATGTCATTTATGCTATGACAGAAA 60  
120 ATGATTTATCTGCTCTCTGCGTTGAGAGATCAAAATGTCATTTATGCTATGACAGAAA 179  
61 ATCTAGAGTGTCCCATCTCTGAGAGTGTCAAGAACCTGTCTCCAAAGTGTGAC 120  
180 ATCTAGAGTGTCCCATCTCTGAGAGTGTCAAGAACCTGTCTCCAAAGTGTGAC 239  
121 CACATATTTTGCAGATTTTGCATGCTGTAACCTTCTCAACGAGAAAGAGGCTTTCACAG 180  
240 CACATATTTTGCAGATTTTGCATGCTGTAACCTTCTCAACGAGAAAGAGGCTTTCACAG 299  
181 TGTCTTTATGTAGAAATGATATACCAAAAGAGGCTTTCACAGAGTGTGAGT 240  
300 TGTCTTTATGTAGAAATGATATACCAAAAGAGGCTTTCACAGAGTGTGAGT 359  
241 CAACTGTGTGAAGCTATTGAAATCATTTGTCTTTTCAGCTTGCACAGGTTTGAG 300  
360 CAACTGTGTGAAGCTATTGAAATCATTTGTCTTTTCAGCTTGCACAGGTTTGAG 419  
301 TATGCAACAGCTATATTTTGCAGAAAGAGAAATTAACCTCTCTCAACCTTCAAAAGAT 360  
420 TATGCAACAGCTATATTTTGCAGAAAGAGAAATTAACCTCTCTCAACCTTCAAAAGAT 479  
361 GAAGTTTCTATCTCAACCAAGTATGAGGCTACAGAAACCGTCCAAAGAGCTTACAGAGT 420  
480 GAAGTTTCTATCTCAACCAAGTATGAGGCTACAGAAACCGTCCAAAGAGCTTACAGAGT 539  
421 GAACCGGAAATCTCTCTGAGAGAAACCAAGTCTGAGTGTCCAACTCTTCAACCTTGA 480  
540 GAACCGGAAATCTCTCTGAGAGAAACCAAGTCTGAGTGTCCAACTCTTCAACCTTGA 539  
481 ACTGTGAGAACTCTGAGAGCAAAAGAGGAGTATCAACCTTCAAAAGAGCTTGTCTACATT 540  
600 ACTGTGAGAACTCTGAGAGCAAAAGAGGAGTATCAACCTTCAAAAGAGCTTGTCTACATT 659  
541 GAATTTGGATCTGATTTCTTGAAGATACCGTTAATTAAGCACTTAATTCAGATGGGGA 600  
660 GAATTTGGATCTGATTTCTTGAAGATACCGTTAATTAAGCACTTAATTCAGATGGGGA 719  
601 GATCAAGAAATTTTACAAATCACCCCTCAAGAGAACGAGATGAATTCAGTTTGATTC 660

|||||  
720 GATCAAGAAATTTTACAAATCACCCCTCAAGAGAACGAGATGAATTCAGTTTGATTC 779  
661 GCAAAAAAGGCTGCTTGTGATTTTTCGAGAGCGAGTATCAAAATTAAGCACTTAATTCAG 720  
780 GCAAAAAAGGCTGCTTGTGATTTTTCGAGAGCGAGTATCAAAATTAAGCACTTAATTCAG 839  
721 CCCAGTATATGATTTTGAACACACACAGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAG 780  
840 CCCAGTATATGATTTTGAACACACACAGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAG 899  
781 TATCAGGAGTGTGATTTTGAACACACACAGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAG 840  
900 TATCAGGAGTGTGATTTTGAACACACACAGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAG 959  
841 AGCTCATTAACAGATGAGAGCAGAGCTTTTATCTCTCAACGAGCAAAATTAAGTGTGAA 900  
960 AGCTCATTAACAGATGAGAGCAGAGCTTTTATCTCTCAACGAGCAAAATTAAGTGTGAA 1019  
901 AAGGCTGAATTCGTATTAATAAGCAACAGAGCTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTG 960  
1020 AAGGCTGAATTCGTATTAATAAGCAACAGAGCTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTG 1079  
961 TGGGCTGGAAGTAAAGAAACATGATATGATAGGCGGAGCTCCAGCAGAGAGAGAGAGAGTAA 1020  
1080 TGGGCTGGAAGTAAAGAAACATGATATGATAGGCGGAGCTCCAGCAGAGAGAGAGAGTAA 1139  
1021 GATCTGAATGCTATCCCTGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080  
1140 GATCTGAATGCTATCCCTGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1199  
1081 TCAGAGATCTCTGAGATGAGATGAGAGAGTGTCTGTGATCACTAATAATGAGCAGATTCAG 1140  
1200 TCAGAGATCTCTGAGATGAGATGAGAGAGTGTCTGTGATCACTAATAATGAGCAGATTCAG 1259  
1141 AAGGTTAATGAGTGTGTTTTCAGAGAGTGAAGTGTGATGTTAGTGTGATGATGATGATGAT 1200  
1260 AAGGTTAATGAGTGTGTTTTCAGAGAGTGAAGTGTGATGTTAGTGTGATGATGATGATGAT 1319  
1201 GGGGAGCTCAATCAAAATGCGCAAGTGTGATGATGATGATGATGATGATGATGATGATGAT 1260  
1320 GGGGAGCTCAATCAAAATGCGCAAGTGTGATGATGATGATGATGATGATGATGATGATGAT 1379  
1261 GAATATCTGCTTCTTACAGAGAAATGAGCTTACGAGGAGTGTGATGATGATGATGATGATG 1320  
1380 GAATATCTGCTTCTTACAGAGAAATGAGCTTACGAGGAGTGTGATGATGATGATGATGATG 1439  
1321 ATATGTAAGTGAAGAGTTCACCTCCAAATCAGTATGAGATATTTGAAGACAAATA 1380  
1440 ATATGTAAGTGAAGAGTTCACCTCCAAATCAGTATGAGATATTTGAAGACAAATA 1499  
1381 TTTGGGAAAACTATGCGAGAGAGGCAAGCTCCCAACTTAAGCCATGTAAGTGAAT 1440  
1500 TTTGGGAAAACTATGCGAGAGAGGCAAGCTCCCAACTTAAGCCATGTAAGTGAAT 1559  
1441 CTATATTAAGAGCATTTGTTACTGAGCCACAGTATTAAGAGGCTCCCTCACAAT 1500  
1560 CTATATTAAGAGCATTTGTTACTGAGCCACAGTATTAAGAGGCTCCCTCACAAT 1619  
1501 AAATTAAGGCTTAAGAGAGAGCTTACATCAGGCTTCAATCAGATTTTATCAAGAAA 1560  
1620 AAATTAAGGCTTAAGAGAGAGCTTACATCAGGCTTCAATCAGATTTTATCAAGAAA 1679  
1561 GCAGATTTGGCAGTTCAAAAGAGCTCCTGAATATTAATGAGGAACTAACCAAGGAG 1620  
1680 GCAGATTTGGCAGTTCAAAAGAGCTCCTGAATATTAATGAGGAACTAACCAAGGAG 1739  
1621 CAGAATGCTCAAGTGAATATTTACTAATGAGGAGTGAATTAATTAAGAGGAT 1680  
1740 CAGAATGCTCAAGTGAATATTTACTAATGAGGAGTGAATTAATTAAGAGGAT 1799  
1681 TCTATTCAAGATGAGAGAAATCCCTAACCCCAATGAAATCACTCAAGAAAGATCTGCTTC 1740  
|||||



Db 1800 TCTATTGAGATGAGAAAAATCTTAACCAATAGAACTACTCGAAAAAGATCTGCTTC 1859  
Qy 1741 AAAAGCAAGCTGACCTATTAAGCAGAGTATTAAGCAATATGCAATCGAATTAATATTC 1800  
Db 1860 AAAAGCAAGCTGACCTATTAAGCAGAGTATTAAGCAATATGCAATCGAATTAATATTC 1919  
Qy 1801 CACAAATTCAAAAGCAGCTTAAGAAATAGGCTGAGAGAAAGTCTTCTACAGGCAATAT 1860  
Db 1920 CCAATATTCAAAAGCAGCTTAAGAAATAGGCTGAGAGAAAGTCTTCTACAGGCAATAT 1979  
Qy 1861 CATGGCCCTTGAAGTACTAGTACAGTAAATCTAAGCCCACTAATTTACTGAATTCGA 1920  
Db 1980 CAGCCCTTGAAGTACTAGTACAGTAAATCTAAGCCCACTAATTTACTGAATTCGA 2039  
Qy 1921 ATTGATAGTGTCTCTAGCAGTGAAGATTAAGAAAAAGTCAACCAATGCTCCAGTC 1980  
Db 2040 ATTGATAGTGTCTCTAGCAGTGAAGATTAAGAAAAAGTCAACCAATGCTCCAGTC 2099  
Qy 1981 AGGCACAGAGAAACCTTACAACTCATGAGAGTAAAGAACCTGCAACTGGAGCCAGAG 2040  
Db 2100 AGGCACAGAGAAACCTTACAACTCATGAGAGTAAAGAACCTGCAACTGGAGCCAGAG 2159  
Qy 2041 AGTAACAGCCCAATATGAAGAGAGTAAAGAAAGATGACAGTACTTTCCAGAGCTG 2100  
Db 2160 AGTAACAGCCCAATATGAAGAGAGTAAAGAAAGATGACAGTACTTTCCAGAGCTG 2219  
Qy 2101 AAGTTAACAATGACCTGCTGTTCTTTACTAAGTGTCAATACAGTGAATTAAGAA 2160  
Db 2220 AAGTTAACAATGACCTGCTGTTCTTTACTAAGTGTCAATACAGTGAATTAAGAA 2279  
Qy 2161 TTTTGTCAATCTAGCTTCCAGAGAGAAAAAGAAAGTGAAGACGTTAAAGTG 2220  
Db 2280 TTTTGTCAATCTAGCTTCCAGAGAGAAAAAGAAAGTGAAGACGTTAAAGTG 2339  
Qy 2221 TCTAATTAATGCTGAAGAACCCCAAGATCTCATGTTAAGTGAAGAGGTTTGGCAACT 2280  
Db 2340 TCTAATTAATGCTGAAGAACCCCAAGATCTCATGTTAAGTGAAGAGGTTTGGCAACT 2399  
Qy 2281 GAAAGATCTGAGAGATGACATATTTCTATTGTAAGTGTGTAAGTGTGACACTGAG 2340  
Db 2400 GAAAGATCTGAGAGATGACATATTTCTATTGTAAGTGTGTAAGTGTGTAAGTGTG 2459  
Qy 2341 GAAAGATCTGAGAGATGACATATTTCTATTGTAAGTGTGTAAGTGTGTAAGTGTG 2400  
Db 2460 GAAAGATCTGAGAGATGACATATTTCTATTGTAAGTGTGTAAGTGTGTAAGTGTG 2519  
Qy 2401 TGGTAGTCTGAGTGCAGATTTGAAACCCCAAGGACTAATTCATGTTGTTCCAA 2460  
Db 2520 TGGTAGTCTGAGTGCAGATTTGAAACCCCAAGGACTAATTCATGTTGTTCCAA 2579  
Qy 2461 GATTAATGAAATGACACAGAGGCTTTAAGTATCCATTGGGACATGAAGTTAACACAGT 2520  
Db 2580 GATTAATGAAATGACACAGAGGCTTTAAGTATCCATTGGGACATGAAGTTAACACAGT 2639  
Qy 2521 CGGGAACAGAGATGAAATGGAAGAAAGTGAAGTGTGATGCTCAGTATTTGCAATACA 2580  
Db 2640 CGGGAACAGAGATGAAATGGAAGAAAGTGAAGTGTGATGCTCAGTATTTGCAATACA 2699  
Qy 2581 TTCAAGGTTTCAAGAGGCGCAGATTTGCTGTTTCAATCCAGAGAAATGGAAGAG 2640  
Db 2700 TTCAAGGTTTCAAGAGGCGCAGATTTGCTGTTTCAATCCAGAGAAATGGAAGAG 2759  
Qy 2641 GAATGTGCAACATTTCTGCCCCACTGCGGTCTTAAAGAAACAAAGTCCAAAAGTACT 2700  
Db 2760 GAATGTGCAACATTTCTGCCCCACTGCGGTCTTAAAGAAACAAAGTCCAAAAGTACT 2819  
Qy 2701 TTTGAATGTGAACAAAAGAGAAATCAAGAGAAAGTGAAGTGTATATCAAGCCTGTA 2760  
Db 2820 TTTGAATGTGAACAAAAGAGAAATCAAGAGAAAGTGAAGTGTATATCAAGCCTGTA 2879  
Qy 2761 CAGACAGTTAATATCTGAGGCTTCTGCTGTGTGTGTCAGAAAGATTAAGCCAGTTGAT 2820  
Db 2880 CAGACAGTTAATATCTGAGGCTTCTGCTGTGTGTGTCAGAAAGATTAAGCCAGTTGAT 2939  
Qy 2821 AATGCCAAATGTAGTATCAAGAGAGGCTGAGGTTTGTCTATCATCTCAGTTGAGAGC 2880  
Db 2940 AATGCCAAATGTAGTATCAAGAGAGGCTGAGGTTTGTCTATCATCTCAGTTGAGAGC 2999  
Qy 2881 AACGAAATGAGTCTATTTACTGCAAAATTAACATGACCTTTTACAAACCCATATGCTATA 2940  
Db 3000 AACGAAATGAGTCTATTTACTGCAAAATTAACATGACCTTTTACAAACCCATATGCTATA 3059  
Qy 2941 CCAGCCCTTTTCCCATCAAGTCAATTTGTTAAACCTAATGTAAAGAAATGCTGACAG 3000  
Db 3060 CCAGCCCTTTTCCCATCAAGTCAATTTGTTAAACCTAATGTAAAGAAATGCTGACAG 3119  
Qy 3001 GAAAGCTTTGAGAGAAATTCATATGTCACCGAAGAAAGTGAAGTGAAGTGAAGTGA 3060  
Db 3120 GAAAGCTTTGAGAGAAATTCATATGTCACCGAAGAAAGTGAAGTGAAGTGAAGTGA 3179  
Qy 3061 AGTAAGTGAAGCAATTTAGCCGTAATTAACATTAAGAGAAATGTTTAAAGAGCCAGC 3120  
Db 3180 AGTAAGTGAAGCAATTTAGCCGTAATTAACATTAAGAGAAATGTTTAAAGAGCCAGC 3239  
Qy 3121 TCAAGCAATTTAATGAGTGAAGTGTGACAGTCAATGAAGTGGGCTCCAGTATTAAGAA 3180  
Db 3240 TCAAGCAATTTAATGAGTGAAGTGTGACAGTCAATGAAGTGGGCTCCAGTATTAAGAA 3299  
Qy 3181 ATAGGTTCCAGTATGAAAAACATTTCAAGCAGAACTAGTGAAGAGAGAGCCAAATTTG 3240  
Db 3300 ATAGGTTCCAGTATGAAAAACATTTCAAGCAGAACTAGTGAAGAGAGAGCCAAATTTG 3359  
Qy 3241 AATGCTATGCTATGATTAAGGGGTTTTCAGAACTGAGTCTATTAACAAAGTCTTCTGGA 3300  
Db 3360 AATGCTATGCTATGATTAAGGGGTTTTCAGAACTGAGTCTATTAACAAAGTCTTCTGGA 3419  
Qy 3301 AGTAATTTAGTATGATCTGTAATAAAGCAAGAAATTAAGAGTATGCTGAGTCTG 3360  
Db 3420 AGTAATTTAGTATGATCTGTAATAAAGCAAGAAATTAAGAGTATGCTGAGTCTG 3479  
Qy 3361 AATAGAGATTTCTCTCATATGATGATTTGATTAAGTGAAGAGGCTATGGAAGTGT 3420  
Db 3480 AATAGAGATTTCTCTCATATGATGATTTGATTAAGTGAAGAGGCTATGGAAGTGT 3539  
Qy 3421 CATGATCTGAGTGTGCTGAGACACTGATGACCTGTTAAGATGATGATGATGATGAT 3480  
Db 3540 CATGATCTGAGTGTGCTGAGACACTGATGACCTGTTAAGATGATGATGATGATGAT 3599  
Qy 3481 GAAGATACTAGTGTGCTGTAATAAGTGAAGAAAGTGTGCTGTTTGAAGAAAGC 3540  
Db 3600 GAAGATACTAGTGTGCTGTAATAAGTGAAGAAAGTGTGCTGTTTGAAGAAAGC 3659  
Qy 3541 GTCCAGAAAGAGAGGCTTGAAGAGAGTCTAGCCCTTTCACCATTAACATTTGGCTGAG 3600  
Db 3660 GTCCAGAAAGAGAGGCTTGAAGAGAGTCTAGCCCTTTCACCATTAACATTTGGCTGAG 3719  
Qy 3601 GGTACCGAAGAGGCGCAAGAAATTAAGTCTCAGAAAGAGAGTATTAAGTGAAGT 3660  
Db 3720 GGTACCGAAGAGGCGCAAGAAATTAAGTCTCAGAAAGAGAGTATTAAGTGAAGT 3779  
Qy 3661 GAAGAGCTTCCGCTTCCACACCTGTTATTTGTAAGTGAAGTGAAGTGAAGTGAAGT 3720  
Db 3780 GAAGAGCTTCCGCTTCCACACCTGTTATTTGTAAGTGAAGTGAAGTGAAGTGAAGT 3839  
Qy 3721 TCTACTAGAGATAGACCCGTTGCTACCGAGTCTGCTGTAAGAAACAGAGAGAGATTTA 3780  
Db 3840 TCTACTAGAGATAGACCCGTTGCTACCGAGTCTGCTGTAAGAAACAGAGAGAGATTTA 3899  
Qy 3781 TTATCATTTAAGAAATAGCTTAAATGACCTGACATACAGATTAATTTGCAAGAGCATCT 3840  
Db 3900 TTATCATTTAAGAAATAGCTTAAATGACCTGACATACAGATTAATTTGCAAGAGCATCT 3959  
Qy 3841 CAGGACATCACCTTGAAGAGAGAAACAAATGCTGCTGATGCTGTTTCTTCCAGAGTC 3900  
Db 3960 CAGGACATCACCTTGAAGAGAGAAACAAATGCTGCTGATGCTGTTTCTTCCAGAGTC 4019



3901 AGTGAATGGAGACTGTGCTGCAAAATACAAACACCAGATCCTTTCTTGATTGTTCT 3960  
4020 AGTGAATGGAGACTGTGCTGCAAAATACAAACACCAGATCCTTTCTTGATTGTTCT 4079  
3961 TCCAAACAAATGAGGATCAGTGTGAAAGCCAGGAGTTGGTGTGATGACAAAGATTG 4020  
4080 TCCAAACAAATGAGGATCAGTGTGAAAGCCAGGAGTTGGTGTGATGACAAAGATTG 4139  
4021 GTTTCAGATGATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4080  
4140 GTTTCAGATGATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4139  
4081 ATGATTCAAACTTAGTGTGAGAGAGAGATCTGGGTGTGAGAGTGAAGAGAGAGAG 4140  
4200 ATGATTCAAACTTAGTGTGAGAGAGAGATCTGGGTGTGAGAGTGAAGAGAGAGAG 4259  
4141 GATGCTCAGGAGTATCTCTCAGAGTGCATTTTAACCACTGACGAGAGAGAGATCATG 4200  
4260 GATGCTCAGGAGTATCTCTCAGAGTGCATTTTAACCACTGACGAGAGAGAGATCATG 4319  
4201 CAACATTAACCTGATTAAGCTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4260  
4320 CAACATTAACCTGATTAAGCTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4379  
4261 CATGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4320  
4380 CATGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4439  
4321 GACTCGGAATCCAGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4380  
4440 GACTCGGAATCCAGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4499  
4381 AGTGAATGGAGACTGTGCTGCAAAATACAAACACCAGATCCTTTCTTGATTGTTCT 4440  
4500 AGTGAATGGAGACTGTGCTGCAAAATACAAACACCAGATCCTTTCTTGATTGTTCT 4559  
4441 GCAGATGTTCTACAGTAAATTAAGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4500  
4560 GCAGATGTTCTACAGTAAATTAAGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4619  
4501 TGCCCATCATTAAGTGAATGAGTGTGATGACAGAGTGTGAGAGAGAGAGAGAGAGAG 4560  
4620 TGCCCATCATTAAGTGAATGAGTGTGATGACAGAGTGTGAGAGAGAGAGAGAGAGAG 4679  
4561 AACTACCATCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4620  
4680 AACTACCATCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4739  
4621 GAGTCTGGGCGACAGATTTGACGGAAATCTTACTTGGCAAGGAGAGAGAGAGAGAG 4680  
4740 GAGTCTGGGCGACAGATTTGACGGAAATCTTACTTGGCAAGGAGAGAGAGAGAGAG 4799  
4681 ACCCTTACTGGAATGGAATCAGCTCTTCTCTATGACCTGGAATGTGATCTTCT 4740  
4800 ACCCTTACTGGAATGGAATCAGCTCTTCTCTATGACCTGGAATGTGATCTTCT 4859  
4741 GAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4800  
4860 GAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4919  
4801 AAAGTTCCCAATTTGAAGTTGAGATCTGCCAGAGAGAGAGAGAGAGAGAGAGAGAG 4860  
4920 AAAGTTCCCAATTTGAAGTTGAGATCTGCCAGAGAGAGAGAGAGAGAGAGAGAGAG 4979  
4861 GATACGCTGGGTATATGCAATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4920  
4980 GATACGCTGGGTATATGCAATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5039  
4921 GCTTCAACAGAAAGGCTCAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4980  
5040 GCTTCAACAGAAAGGCTCAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5099  
4981 GAATTTATGCTCGTGTACAGAGTTGGCAGAAACACCAATCACTTAATTAAT 5040

5100 GAAATTTATGCTCGTGTACAGAGTTGGCAGAAACACCAATCACTTAATTAATTAAT 5159  
5041 ACTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5100  
5160 ACTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5219  
5101 CTGAATATTTTCTAGAGATTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5160  
5220 CTGAATATTTTCTAGAGATTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5279  
5161 CAGCTATTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5220  
5280 CAGCTATTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5339  
5221 GTCAATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5280  
5340 GTCAATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5399  
5281 TTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5340  
5400 TTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5459  
5341 GAATGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5400  
5460 GAATGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5519  
5401 GGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5460  
5520 GGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5579  
5461 TTCCATGCAATTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5520  
5580 TTCCATGCAATTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5579  
5521 AGTGTACACTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5580  
5640 AGTGTACACTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5699  
5581 AGCCACTAC 5589  
5700 AGCCACTAC 5708

RESULT 7  
US-09-982-828-1  
: Sequence 1, Application US/09982828  
: Publication No. US20030022184A1  
: GENERAL INFORMATION:  
: APPLICANT: Murphy, Patricia D.  
: Allen, Antonette C.  
: Alvarez, Christopher P.  
: Critz, Brenda S.  
: Olson, Sheri J.  
: Thurber, Denise  
: Zeng, Bin  
: TITLE OF INVENTION: Coding Sequences of the Human  
: NUMBER OF SEQUENCES: 72  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Morgan Lewis & Bockius LLP  
: STREET: 1111 Pennsylvania Avenue N. W.  
: CITY: Washington  
: STATE: DC  
: COUNTRY: USA  
: ZIP: 20004  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: PatentIn Release #1.0, Version #1.30  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/09/982,828

FILED DATE: 22-Oct-2001  
CLASSIFICATION: <unknown>  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 09/074,453  
FILING DATE: 1998-05-06  
APPLICATION NUMBER: US 08/798,691  
FILING DATE: 1997-02-12  
APPLICATION NUMBER: US 08/598,591  
FILING DATE: 1996-02-12  
ATTORNEY/AGENT INFORMATION:  
NAME: Michael S. Tuscan  
REGISTRATION NUMBER: 43,210  
REFERENCE/DOCKET NUMBER: 44921-5053-01-05  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-739-3000  
TELEFAX: 202-739-3001  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5711 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
STRAIN: BRCA1 (om11)  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: 17  
MAP POSITION: 17q21  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-982-828-1

Query Match 99.9%; Score 5581; DB 9; Length 5711;  
Best Local Similarity 99.9%; Pred. No. 0;

Matches 5584; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

1 ATGAGATTATCTGCTCTTGGCTTGAGAGAAAGTCAATTAATGCTATGACAAA 60  
120 ATGAGATTATCTGCTCTTGGCTTGAGAGAAAGTCAATTAATGCTATGACAAA 179  
61 ATCTGAGATGTCCTGCTGCTGAGAGTGTATCAAGAACCTGCTCCACAAAGTGAC 120  
180 ATCTGAGATGTCCTGCTGCTGAGAGTGTATCAAGAACCTGCTCCACAAAGTGAC 239  
121 CACATATTTGGCAATTTTGCATGTGAACCTTCTCAACGAGAAAGGCGCTTCACAG 180  
240 CACATATTTGGCAATTTTGCATGTGAACCTTCTCAACGAGAAAGGCGCTTCACAG 299  
181 TGTCTTTATGTAGATGATATACCAAAAGAGCCTACAGAAAGTACGAGATTAGT 240  
300 TGTCTTTATGTAGATGATATACCAAAAGAGCCTACAGAAAGTACGAGATTAGT 359  
241 CAACCTGTTGAAGAGCTATGAAATCATTTGCTTTTCAAGCTTGACAGGTTGGAG 300  
360 CAACCTGTTGAAGAGCTATGAAATCATTTGCTTTTCAAGCTTGACAGGTTGGAG 419  
301 TATGCAAAAGAGCTATGAAATCATTTGCTTTTCAAGCTTGACAGGTTGGAG 360  
420 TATGCAAAAGAGCTATGAAATCATTTGCTTTTCAAGCTTGACAGGTTGGAG 479  
361 GAAGTTTCTATCATCCAAAGATAGGCTACAGAAACCGCTGCCAAAAGACTTCTACAGAGT 420  
480 GAAGTTTCTATCATCCAAAGATAGGCTACAGAAACCGCTGCCAAAAGACTTCTACAGAGT 539  
421 GAACCGGAAATCTCTTGGAGAAACAGCTCAGTGTCCAGCTCTTAACCTTGA 480  
540 GAACCGGAAATCTCTTGGAGAAACAGCTCAGTGTCCAGCTCTTAACCTTGA 599  
481 ACTGTGAGAACTCTGAGCAAGAGAGGATACAACTCAAAAGAGCTGTGCTACATT 540  
600 ACTGTGAGAACTCTGAGCAAGAGAGGATACAACTCAAAAGAGCTGTGCTACATT 659  
541 GAATTGGATCTGATTTCTTCAAGATACCGTTAATTAAGCAACTTAATTCAGTGGGA 600

660 GAATTGGATCTGATTTCTTCAAGATACCGTTAATTAAGCAACTTAATTCAGTGGGA 719  
601 GATCAAGAAATTTTACAAATCACCCTCAAGAGACCGAGGAAATCGTTGATCT 660  
720 GATCAAGAAATTTTACAAATCACCCTCAAGAGACCGAGGAAATCGTTGATCT 779  
661 GCAAAAAGAGCTCTTGTGAATTTTGTAGAGAGGATGTAACAATACTGACATCA 720  
780 GCAAAAAGAGCTCTTGTGAATTTTGTAGAGAGGATGTAACAATACTGACATCA 839  
721 CCCAGTATATATGATTTGAACACCACTGAGAGCGTGCAGCTGAGAGCATCCAGAAAG 780  
840 CCCAGTATATATGATTTGAACACCACTGAGAGCGTGCAGCTGAGAGCATCCAGAAAG 899  
781 TATCAGGATGCTCTGTTCAAACTGCAATGGAGCCATGGGCAACAATCTCATGCC 840  
900 TATCAGGATGCTCTGTTCAAACTGCAATGGAGCCATGGGCAACAATCTCATGCC 959  
841 AGCTCATTTACAGATGAGAAACAGCACTTATTACTACTAAGACAGAAATGATAGAA 900  
960 AGCTCATTTACAGATGAGAAACAGCACTTATTACTACTAAGACAGAAATGATAGAA 1019  
901 AAGCTGATTTCTGTAATTAATAAGCAACAGCTGCTTACGAAAGAGCCACATTAACGA 960  
1020 AAGCTGATTTCTGTAATTAATAAGCAACAGCTGCTTACGAAAGAGCCACATTAACGA 1079  
961 TGGGCTGGAAGTAAAGGAAACATGTAATGATGAGCGGACCTCCAGCAACAAAGGTA 1020  
1080 TGGGCTGGAAGTAAAGGAAACATGTAATGATGAGCGGACCTCCAGCAACAAAGGTA 1139  
1021 GATCTGAATGCTGATCCCTGTGTGAGAGAAAGATGATAGAGAAATGCTGATGC 1080  
1140 GATCTGAATGCTGATCCCTGTGTGAGAGAAAGATGATAGAGAAATGCTGATGC 1199  
1081 TCAGAGATCTTGAGATGATGAGATGTTCTTGTGATTAACATTAATAGCAGATTAG 1140  
1200 TCAGAGATCTTGAGATGATGAGATGTTCTTGTGATTAACATTAATAGCAGATTAG 1259  
1141 AAAGTTAATGATGTTGTTTCCAGAGATGATGATGTTGTTGATGATGATGATGAT 1200  
1260 AAAGTTAATGATGTTGTTTCCAGAGATGATGATGTTGTTGATGATGATGATGAT 1319  
1201 GGGGAGTCTGAATCAATAGCCAAAGTACGTATGATTTGAGAGTTCTAATGAGGTAGAT 1260  
1320 GGGGAGTCTGAATCAATAGCCAAAGTACGTATGATTTGAGAGTTCTAATGAGGTAGAT 1379  
1261 GAATATTTCTGTTCTTCAAGAAATAGACTTACTGCGCAGATCTCTATGAGCTTTA 1320  
1380 GAATATTTCTGTTCTTCAAGAAATAGACTTACTGCGCAGATCTCTATGAGCTTTA 1439  
1321 ATATGTAAGGTAAGAGATGCTCAATCTGATGAGAGATATTTGAGAGCAAAATA 1380  
1440 ATATGTAAGGTAAGAGATGCTCAATCTGATGAGAGATATTTGAGAGCAAAATA 1499  
1381 TTTGGGAAAACCTATCGGAGAAAGAGCAAGCCTCCCACTTAAGCATGTAACTGAAAT 1440  
1500 TTTGGGAAAACCTATCGGAGAAAGAGCAAGCCTCCCACTTAAGCATGTAACTGAAAT 1559  
1441 CTATATTTAGAGATTTTGTACTGAGCCAGATATATCAAGAGCGTCCCTCAAAAT 1500  
1560 CTATATTTAGAGATTTTGTACTGAGCCAGATATATCAAGAGCGTCCCTCAAAAT 1619  
1501 AAATTTAAGGTTAAAGGAGACCTACATCAAGGCTTCACTGAGAGATTTTATCAAGAA 1560  
1620 AAATTTAAGGTTAAAGGAGACCTACATCAAGGCTTCACTGAGAGATTTTATCAAGAA 1679  
1561 GCAGATTTGGCAGTTCAAAAGACTCTGTAATGATTAATCAGGAACTTAACCAAGCGAG 1620  
1680 GCAGATTTGGCAGTTCAAAAGACTCTGTAATGATTAATCAGGAACTTAACCAAGCGAG 1739  
1621 CAGATGCTCAAGTGAATATTAATTAATGCTCATGAGATTAACCAAGCGAG 1680

QY	2761	CAGCAGAGTTAATATCAGACGACGGCTTCTCGTGGTGGCAGAAAGATTAAGCCACTGAT	282
Db	2880	CAGACAGTTAATATACACTGACAGGCTTCTCGTGGTGGCAGAAAGATTAAGCCACTGAT	293
QY	2821	AATGCCAATATAGTATCAMAAGGAGCGCTAGTGTTCGTATATCATCTACGTTCCAGAGC	2886
Db	2940	AATGCCAATATAGTATCAMAAGGAGCGCTAGTGTTCGTATATCATCTACGTTCCAGAGC	2999
QY	2881	AACGAACTGGACTATTACTTCCAAATTAACATGACCTTTTACAAACCCTATTCGATTA	2940
Db	3000	AACGAACTGGACTATTACTTCCAAATTAACATGACCTTTTACAAACCCTATTCGATTA	3059
QY	2941	CCACACATTTTCCCATCAGATGCATTTGTTTAAACTTAATATGAAAGAAATCTGCTAG	3000
Db	3060	CCACACATTTTCCCATCAGATGCATTTGTTTAAACTTAATATGAAAGAAATCTGCTAG	3119
QY	3001	GAAGAACTTTGAGGAACATTTCAATGTCACCTGAAAGAGAAATGGGAAATGAGAACATTTCCA	3060
Db	3120	GAAGAACTTTGAGGAACATTTCAATGTCACCTGAAAGAGAAATGGGAAATGAGAACATTTCCA	3179
QY	3061	AGTACAGTGCACACATTTAGCCGTATATAACATTAGAGAAATGTTTTTAAAGAACCCAGC	3120
Db	3180	AGTACAGTGCACACATTTAGCCGTATATAACATTAGAGAAATGTTTTTAAAGAACCCAGC	3239
QY	3121	TCACCAATATTTAATGAGTAGAGTGGTCCAGTACTATGTAAGTGGGCTCCAGTATTAATCA	3180
Db	3240	TCACCAATATTTAATGAGTAGAGTGGTCCAGTACTATGTAAGTGGGCTCCAGTATTAATCA	3299
QY	3181	ATAGTCTCCAGTGATGAAGACATTCACAGACACTAGTAGAAGACAGAGGCCAAATTG	3240
Db	3300	ATAGTCTCCAGTGATGAAGACATTCACAGACACTAGTAGAAGACAGAGGCCAAATTG	3359
QY	3241	AATGTATGCTTAAAGTTAGTGGGTTTTGGACCTGAGGCTCTTAAACAAGCTTCCTGGA	3300
Db	3360	AATGTATGCTTAAAGTTAGTGGGTTTTGGACCTGAGGCTCTTAAACAAGCTTCCTGGA	3419
QY	3301	AGTAATTTGTAAGCATCTGGAATAAAAAAGACAGAAATTTGAAGAGTAGTTCCAGACTGTT	3360
Db	3420	AGTAATTTGTAAGCATCTGGAATAAAAAAGACAGAAATTTGAAGAGTAGTTCCAGACTGTT	3479
QY	3361	AATACAGATTTCTCCCATATCTGATTTCAATATACTTAGAACAAGCCCTATGGAAGTAGT	3420
Db	3480	AATACAGATTTCTCCCATATCTGATTTCAATATACTTAGAACAAGCCCTATGGAAGTAGT	3539
QY	3421	CATGATCTCAGGTTTTGTTCTGAGACACCTGATACCTCGTAGATGATGGTAATTAAG	3480
Db	3540	CATGATCTCAGGTTTTGTTCTGAGACACCTGATACCTCGTTAGATGATGGTAATTAAG	3599
QY	3481	GAAGATACTATTTTGCTGAAATGACATTAAGAGAAAGTCTGCTGTTTTAGCAAAAGC	3540
Db	3600	GAAGATACTATTTTGCTGAAATGACATTAAGAGAAAGTCTGCTGTTTTAGCAAAAGC	3659
QY	3541	GTCAGAAAGAGAGAGCTTAGCAGAGAGTCGAGCCCTTACACCCATACACATTTGGCTAG	3600
Db	3660	GTCAGAGAGAGAGAGCTTAGCAGAGAGTCGAGCCCTTACACCCATACACATTTGGCTAG	3719
QY	3601	GGTTACCGAAGAGGGGCCAAGAAATTTAGAGTCTCAGAGAGAACTATCTAGTGAAGAT	3660
Db	3720	GGTTACCGAAGAGGGGCCAAGAAATTTAGAGTCTCAGAGAGAACTATCTAGTGAAGAT	3779
QY	3661	GAAAGAGCTCCCTGCTTCCAAACACTTGTATTTGGTAAAGTAACATATATCTTTCAG	3720
Db	3780	GAAAGAGCTCCCTGCTTCCAAACACTTGTATTTGGTAAAGTAACATATATCTTTCAG	3839
QY	3721	TCTACTAGAGGCTATACACCGTGTGCTACCGAGTGTGCTCTAAGAAACAGAGAGGAATTTA	3780
Db	3840	TCTACTAGAGGCTATACACCGTGTGCTACCGAGTGTGCTCTAAGAAACAGAGAGGAATTTA	3899
QY	3781	TTATCATTTGAAGATAGCTTAAATGACTGCAGTACACAGGTAATATTTGGCAAAAGCATCT	3840
Db	3900	TTATCATTTGAAGATAGCTTAAATGACTGCAGTACACAGGTAATATTTGGCAAAAGCATCT	3959

OY	3841	CAGGAACATCACCTTAAGGAGAAACAAATGTTCTGCTAGCTGTCTTTCTTCCACAGTGC	3900
Db	3960	CAGGACATCCACTTTAGTAGGAAACAAATGTTCTGCTAGCTGTCTTTCTTCCACAGTGC	4019
OY	3901	AGTGAATTTGGAAGACTTACTCTGACATCTCAAAATCCAAACACCACAGATCCCTTCTTGTGATTTGGTCT	3960
Db	4020	AGTGAATTTGGAAGACTTACTCTGACATCTCAAAATCCAAACACCACAGATCCCTTCTTGTGATTTGGTCT	4079
OY	3961	TTCAAACAAATGAGGCATCTCTGAAAGCCAGGAGTTCGTTGAGTGTGACAAAGAAATTTG	4020
Db	4080	TTCAAACAAATGAGGCATCTCTGAAAGCCAGGAGTTCGTTGAGTGTGACAAAGAAATTTG	4139
OY	4021	GTTTGAATGATGAAGAAAGGAAGGAGGCTTTGGAAGAAATTAATCAAGAAAGCAAGC	4080
Db	4140	GTTTGAATGATGAAGAAAGGAAGGAGGCTTTGGAAGAAATTAATCAAGAAAGCAAGC	4199
OY	4081	ATGATTTCAAACTTAGTGTGAAGCAGACATCTGGGTGTGAGAGTGAACCAAGGCTCTGTGAA	4140
Db	4200	ATGATTTCAAACTTAGTGTGAAGCAGACATCTGGGTGTGAGAGTGAACCAAGGCTCTGTGAA	4259
OY	4141	GACTGCTCAGGGCTATCTCTTCAGAGTACATTTTAACTTACAGCAGAGGAGATACATG	4200
Db	4260	GACTGCTCAGGGCTATCTCTTCAGAGTACATTTTAACTTACAGCAGAGGAGATACATG	4319
OY	4201	CAACATACCTGTATTAAGCTCCAGCAGAGAAATGCGTGAACATTAAGAAAGCTGTGTAGAACAG	4260
Db	4320	CAACATACCTGTATTAAGCTCCAGCAGAGAAATGCGTGAACATTAAGAAAGCTGTGTAGAACAG	4379
OY	4261	CATGAGGAGCCAGCCCTTCTAACAGTACCCCTTCATCATTAAGTGAATCTCTTCGCCCTTACG	4320
Db	4380	CATGAGGAGCCAGCCCTTCTAACAGTACCCCTTCATCATTAAGTGAATCTCTTCGCCCTTACG	4439
OY	4321	GACCTGCGAAATTCGAAACAAAGCAGATCAGAAAAAGCAGTAATTAATTCACAGAAAAAGT	4380
Db	4440	GACCTGCGAAATTCGAAACAAAGCAGATCAGAAAAAGCAGTAATTAATTCACAGAAAAAGT	4499
OY	4381	AGTGAATTCCTTATTAAGCCAGAAATTCGAAAGGCTTTGCTGTACAAATTTGAGAGTGTCT	4440
Db	4500	AGTGAATTCCTTATTAAGCCAGAAATTCGAAAGGCTTTGCTGTACAAATTTGAGAGTGTCT	4559
OY	4441	GCAGATAGTCTTACACAGTAAATTAAGAACAGAGAGTGAAGAAAGTCACTCCCTCTCTTAA	4500
Db	4560	GCAGATAGTCTTACACAGTAAATTAAGAACAGAGAGTGAAGAAAGTCACTCCCTCTCTTAA	4619
OY	4501	TGCCCATCTTATGATGATAGTGTGTACATGACAGTGTGCTGTGGAGTCTTCAGAAATAGA	4560
Db	4620	TGCCCATCTTATGATGATAGTGTGTACATGACAGTGTGCTGTGGAGTCTTCAGAAATAGA	4679
OY	4561	AACATACCCATCTCAAGAGAGACTCATTAAGTGTGTATTTGTGAGAGAGCAACACTGGAA	4620
Db	4680	AACATACCCATCTCAAGAGAGACTCATTAAGTGTGTATTTGTGAGAGAGCAACACTGGAA	4739
OY	4621	GAGTCTGGGGCAGACAGATTTGACGGAACATCTTACTTGCCCAAGCAAGATCTAGAGGGA	4680
Db	4740	GAGTCTGGGGCAGACAGATTTGACGGAACATCTTACTTGCCCAAGCAAGATCTAGAGGGA	4799
OY	4681	ACCCCTTAACCTGGAATCTGGAATCAGCCTCTTCTGTGATGACCCCTGATCTGATCTCTCT	4740
Db	4800	ACCCCTTAACCTGGAATCTGGAATCAGCCTCTTCTGTGATGACCCCTGATCTGATCTCTCT	4859
OY	4741	GAAACAGAGCCCCAGAGTCAAGCTGTGTGTGCAACATACCATCTTCAACCTGTGATG	4800
Db	4860	GAAACAGAGCCCCAGAGTCAAGCTGTGTGTGCAACATACCATCTTCAACCTGTGATG	4919
OY	4801	AAAGTTCCTCCCAATTTGAAAGTTGCAAGATCTGCCAGAGTCCAGTGTGCTCATACTACT	4860
Db	4920	AAAGTTCCTCCCAATTTGAAAGTTGCAAGATCTGCCAGAGTCCAGTGTGCTCATACTACT	4979
OY	4861	GATACCTGCTGGGTATTAATGCAATGGAAGAAAGTGTGAGCAGAGGAGAGCCAGAAATTTACA	4920
Db	4980	GATACCTGCTGGGTATTAATGCAATGGAAGAAAGTGTGAGCAGAGGAGAGCCAGAAATTTACA	5039
OY	4921	GCCTTCAACAGAAAGGCTCAACAAAAAGATGCTCATGTGCTGTGCTGACCTGACCCAGAA	4980

Db	5040	GCTTCAACAGAAAGGCTACAAAGAAAGATGTCATGGTGGTGTGCTGGCCGACCCAGAA	5099
OY	4981	GAATTATATGCGCTGCTACAAAGTTTGGCCAGAAACACACATCACTTAATTAATTAAT	5040
Db	5100	GAATTTATGCGCGTGTCACAAAGTTTGGCCAGAAACACACATCACTTAATTAATTAAT	5159
OY	5041	ACTAAGACGACTACTCATGTTGTTATGAAACAGATGCTGAGTTTGTGTGACGGACA	5100
Db	5160	ACTAAGACGACTACTCATGTTGTTATGAAACAGATGCTGAGTTTGTGTGACGGACA	5219
OY	5101	CTGAATATATTTCTAGCAATTTGGCGGGAGAAATGGGTAGTTAGCTATTTGCGGTGACC	5160
Db	5220	CTGAATATATTTCTAGCAATTTGGCGGGAGAAATGGGTAGTTAGCTATTTGCGGTGACC	5279
OY	5161	CAGTCTATTTAAGAAAGAAATAATGCTGTAATGAGCATGATTTTGAAGTCAGAGAGATGTG	5220
Db	5280	CAGTCTATTTAAGAAAGAAATAATGCTGTAATGAGCATGATTTTGAAGTCAGAGAGATGTG	5339
OY	5221	GTCATATGGAAAGAAACCAACCAAGTCCAAAGCCAGCAAGAGATATCCAGACAGAAAGATC	5280
Db	5340	GTCATATGGAAAGAAACCAACCAAGTCCAAAGCCAGCAAGAGATATCCAGACAGAAAGATC	5399
OY	5281	TTTCAGGGGGGTAGAAATCTGTGCTATAGGGCCCTTCACACATATGCCACAGATCAAGT	5340
Db	5400	TTTCAGGGGGGTAGAAATCTGTGCTATAGGGCCCTTCACACATATGCCACAGATCAAGT	5459
OY	5341	GAATGATGTGTACAGCTGTGTGTGCTTCTGTGTGTGAAGAGCTTTCATCATTTCAACCTT	5400
Db	5460	GAATGATGTGTACAGCTGTGTGTGCTTCTGTGTGTGAAGAGCTTTCATCATTTCAACCTT	5519
OY	5401	GGCACAGGTTGCCACCAATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	5460
Db	5520	GGCACAGGTTGCCACCAATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	5579
OY	5461	TTTCATCATTAATTTGGGCGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	5520
Db	5580	TTTCATCATTAATTTGGGCGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	5639
OY	5521	AGTGTAGCACTTACACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	5580
Db	5640	AGTGTAGCACTTACACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	5699
OY	5581	AGGCACCTAC 5589	
Db	5700	AGGCACCTAC 5708	

```

RESULT 8
US-09-911-904-127
; Sequence 127, Application US/09/911904
; Publication No. US20030096234A1
; GENERAL INFORMATION:
; APPLICANT: Farr, Spencer B.
; APPLICANT: Pickett, Gavin G.
; APPLICANT: Neft, Rodin Eileen
; APPLICANT: Dunn, II, Robert Thomas
; TITLE OF INVENTION: CANINE TOXICITY GENES
; FILE REFERENCE: 400742000200
; CURRENT APPLICATION NUMBER: US/09/911,904
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: US 60/220,057
; PRIOR FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 386
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 127
; LENGTH: 499
; TYPE: DNA
; ORGANISM: Canis familiaris
US-09-911-904-127

Query Match          6.5%; Score 364.6; DB 9; Length 499;
Best Local Similarity 86.7%; Pted. No. 4,4e-81;

```

Matches	419	Conservative	0	Mismatches	49	Indels	15	Gaps	1
QY	5110	TTTCTAGAAATTCGGGAGAGAAAATGGGTAGTACTATTTCGTGGGTACCCAGTCTATT							5169
Db	1	TTTCTGGTATTTTCAGAGGAAAATGGGTAGTACTATTTCGTGGGTACCCAGTCTATT							60
QY	5170	AAGAAGAAATATGCTGAAATGACATGATTTTGAAGTCAAGAGAAATGTGTCATATGA							5229
Db	61	AAAGAAAGAAAGATACTAGATGAGCATGATTTTGAAGTCAAGAGAAATGTGTCATATGA							120
QY	5230	AGAAACCAACCAAGGTCCAAAGCGAGACAGAAATCCAG-----GACAGA							5274
Db	121	AGAAATACCAAGGTCCAAAGCGAGACAGAAATCCAGAGACAGAAATCCAAAGACAGA							180
QY	5275	AAATCTTCAGGGGGCTAGAAATCTGTGCTATGAGGCCCTTCACCAATATGCCACAGAT							5334
Db	181	AAATCTTCAGGGGGCTAGAAATCTGTGCTATGAGGCCCTTCACCAATATGCCACAGAT							240
QY	5335	CAACTGGAATGAATGATGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT							5394
Db	241	CAATTTGAGTGGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT							300
QY	5395	ACCCCTGGCGACAGGTGTCCACCCCAATTTGTTGTGTGTGTGTGTGTGTGTGTGTGTGT							5454
Db	301	ACCCCTGACGAGAGGGGACACTATCCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT							360
QY	5455	AATGCTTCCATCAATTTGGGAGATGTGTGAGGACCTGTGTGTGTGTGTGTGTGTGTGTGT							5514
Db	361	AGTGCTTCCATCAATTTGGGAGATGTGTGAGGACCTGTGTGTGTGTGTGTGTGTGTGTGT							420
QY	5515	TTTGACAGTGTACACTTCTTACCAAGTGTGACAGAGCTGTGAACTTACTGTATACCCAGAT							5574
Db	421	CTGAGACAGTGTACACTTCTTACCAAGTGTGACAGAGCTGTGAACTTACTGTATACCCAGAT							480
QY	5575	CCC 5577							
Db	481	CCC 483							
RESULT 9									
US-09-864-761-4552									
Sequence 4552, Application US/09864761									
Patent No. US20020048763A1									
GENERAL INFORMATION:									
APPLICANT: Penn, Sharon G.									
APPLICANT: Rank, David R.									
APPLICANT: Hanzel, David K.									
APPLICANT: Chen, Wensheng									
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR									
FILE REFERENCE: Aeomica-X-1									
CURRENT APPLICATION NUMBER: US/09/864,761									
CURRENT FILING DATE: 2001-05-23									
PRIORITY APPLICATION NUMBER: US 60/180,312									
PRIORITY FILING DATE: 2000-02-04									
PRIORITY APPLICATION NUMBER: US 60/207,456									
PRIORITY FILING DATE: 2000-05-26									
PRIORITY APPLICATION NUMBER: US 09/632,366									
PRIORITY FILING DATE: 2000-08-03									
PRIORITY APPLICATION NUMBER: GB 24263.6									
PRIORITY FILING DATE: 2000-10-04									

```

PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
SEQ ID NO 4552
LENGTH: 424
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO L78833.1
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.4
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.1
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.3
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.9
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.4
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.3
US-09-864-761-4552

Query Match      3.1%; Score 175.4; DB 10; Length 424;
Best Local Similarity 94.3%; Pred. No. 1.1e-33;
Matches 182; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY      4184 AGCAGAGGGAATACCATCAACATTAACCTGTATAAGCTCCAGCAGAAATGGCTGAAGTAG 4243
        |||||
Db       232 AGCAGAGGGAATACCATCAACATTAACCTGTATAAGCTCCAGCAGAAATGGCTGAAGTAG 291

QY      4244 AAGCGTGTAGAAGAACACATGGGAGGCCAGCTTCATACACTCCCTTCATCATATAGTG 4303
        |||||
Db       292 AAGCTGTGTAGAAGAACACATGGGAGGCCAGCTTCATACACTCCCTTCATCATATAGTG 351

QY      4304 ACTCTTCTGCCTTTGAGAGACCTGCGAATATCCAGAACAAACACATCAGAAAAAGCAGTAT 4363
        |||||
Db       352 ACCTCTTCTGCCTTTGAGAGACCTGCGAATATCCAGAACAAACACATCAGAAAAAGGTCGT 411

QY      4364 TAACTTCACAGAA 4376
        || |
Db       412 ATTGTGGCCAAA 424

RESULT 10
US-09-864-761-21299
Sequence 21299, Application US/09864761
Patent No. US20020048763a1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aeomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04

```

PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 21299  
LENGTH: 147  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO L78833.1  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.4  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.1  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.3  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.9  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.4  
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.3  
OTHER INFORMATION: EST HUMAN HIT: AUI25312.1, EVALUE 2.00e-77  
OTHER INFORMATION: SWISSPROT HIT: P38398, EVALUE 4.00e-12  
OTHER INFORMATION: NT HIT: g111424969, EVALUE 1.00e-77  
US-09-864-761-21299

Query Match 2.6%; Score 147; DB 10; Length 147;  
Best Local Similarity 100.0%; Pred. No. 7.6e-27;  
Matches 147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 4211 TGATAAAGCTCCAGCAGGAATGCTGAAGTGTGTTAGAACATGGAGCC 4270  
1 TGATAAAGCTCCAGCAGGAATGCTGAAGTGTGTTAGAACATGGAGCC 60  
QY 4211 AGCCTTTACAGCTACCCCTTCATCATATAGTACTCTTCCCTTGAGACCTGGCAA 4330  
|||||  
DB 61 AGCCTTTACAGCTACCCCTTCATCATATAGTACTCTTCCCTTGAGACCTGGCAA 120  
4331 ATCCGAAACAAGCATCAGAAAAG 4357  
|||||  
DB 121 ATCCGAAACAAGCATCAGAAAAG 147

RESULT 11  
US-09-818-875-661  
Sequence 661, Application US/09818875  
Publication No. US20030051270A1  
GENERAL INFORMATION:  
APPLICANT: Kmiec, Eric B.  
APPLICANT: Gamper, Howard B.  
TITLE OF INVENTION: Targeted Chromosomal Genomic Alterations with Modified Single  
TITLE OF INVENTION: Stranded Oligonucleotides  
FILE REFERENCE: Napro-4  
CURRENT APPLICATION NUMBER: US/09/818,875  
CURRENT FILING DATE: 2001-03-27  
PRIOR APPLICATION NUMBER: US 60/192,176  
PRIOR FILING DATE: 2000-03-27  
PRIOR APPLICATION NUMBER: US 60/192,179  
PRIOR FILING DATE: 2000-03-27  
PRIOR APPLICATION NUMBER: US 60/208,538  
PRIOR FILING DATE: 2000-06-01  
PRIOR APPLICATION NUMBER: US 60/244,989  
PRIOR FILING DATE: 2000-10-30  
NUMBER OF SEQ ID NOS: 4385  
SOFTWARE: Friedman macro Napro4  
SEQ ID NO 661  
LENGTH: 121  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-818-875-661

Query Match 2.2%; Score 121; DB 9; Length 121;  
Best Local Similarity 100.0%; Pred. No. 2.2e-20;  
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGATTATATCTGCTCTTCCCTTGAGAGTACAAATGCTATATGCTATCCAGAA 60  
|||||  
DB 1 ATGATTATATCTGCTCTTCCCTTGAGAGTACAAATGCTATATGCTATCCAGAA 60  
61 ATCTTAGAGTGTCCATCTGTCTGAGTTGATCAAGAACCTGTCTCCACAAAGTGAC 120  
QY 121 C 121  
DB 121 C 121

RESULT 12  
US-09-818-875-662/c  
Sequence 662, Application US/09818875  
Publication No. US20030051270A1  
GENERAL INFORMATION:  
APPLICANT: Kmiec, Eric B.  
APPLICANT: Gamper, Howard B.  
TITLE OF INVENTION: Targeted Chromosomal Genomic Alterations with Modified Single  
TITLE OF INVENTION: Stranded Oligonucleotides  
FILE REFERENCE: Napro-4  
CURRENT APPLICATION NUMBER: US/09/818,875  
CURRENT FILING DATE: 2001-03-27  
PRIOR APPLICATION NUMBER: US 60/192,176  
PRIOR FILING DATE: 2000-03-27  
PRIOR APPLICATION NUMBER: US 60/192,179  
PRIOR FILING DATE: 2000-03-27  
PRIOR APPLICATION NUMBER: US 60/208,538  
PRIOR FILING DATE: 2000-06-01  
PRIOR APPLICATION NUMBER: US 60/244,989  
PRIOR FILING DATE: 2000-10-30  
NUMBER OF SEQ ID NOS: 4385  
SOFTWARE: Friedman macro Napro4  
SEQ ID NO 662  
LENGTH: 121  
TYPE: DNA

```

: ORGANISM: Homo sapiens
US-09-818-875-662

Query Match
Best Local Similarity 100.0%; Score 121; DB 9; Length 121;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAATTATCTGCTCTCCGTTGAAGAAGTACAAATGCTATATGCTATGACAGAAA 60
Db 121 ATGGAATTATCTGCTCTCCGTTGAAGAAGTACAAATGCTATATGCTATGACAGAAA 62

QY 61 ATCTTAGAGTGTCCCATCTGCTGAGAGTTGATCAAGAACCTGTCTCCACAAAGTGTGAC 120
Db 61 ATCTTAGAGTGTCCCATCTGCTGAGAGTTGATCAAGAACCTGTCTCCACAAAGTGTGAC 2

QY 121 C 121
Db 1 C 1

RESULT 13
US-09-818-875-665
: Sequence 665, Application US/09818875
: Publication No. US20030051270A1
: GENERAL INFORMATION:
: APPLICANT: Kmiec, Eric B.
: APPLICANT: Gamper, Howard B.
: APPLICANT: Rice, Michael C.
: TITLE OF INVENTION: Targeted Chromosomal Genomic Alterations with Modified Single
: TITLE OF INVENTION: Stranded Oligonucleotides
: FILE REFERENCE: Napro-4
: CURRENT APPLICATION NUMBER: US/09/818,875
: CURRENT FILING DATE: 2001-03-27
: PRIOR APPLICATION NUMBER: US 60/192,176
: PRIOR FILING DATE: 2000-03-27
: PRIOR APPLICATION NUMBER: US 60/192,179
: PRIOR FILING DATE: 2000-03-27
: PRIOR APPLICATION NUMBER: US 60/208,538
: PRIOR FILING DATE: 2000-06-01
: PRIOR APPLICATION NUMBER: US 60/244,989
: PRIOR FILING DATE: 2000-10-30
: NUMBER OF SEQ ID NOS: 4385
: SOFTWARE: Friedman macro Napro4
: SEQ ID NO 665
: LENGTH: 121
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-818-875-665

Query Match
Best Local Similarity 100.0%; Score 121; DB 9; Length 121;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 ATTATATCTGCTCTCCGTTGAAGAAGTACAAATGCTATATGCTATGACAGAAATCT 64
Db 1 ATTATATCTGCTCTCCGTTGAAGAAGTACAAATGCTATATGCTATGACAGAAATCT 60

QY 65 TAGAGTGTCCCATCTGCTGAGAGTTGATCAAGAACCTGTCTCCACAAAGTGTGACACA 124
Db 61 TAGAGTGTCCCATCTGCTGAGAGTTGATCAAGAACCTGTCTCCACAAAGTGTGACACA 120

QY 125 T 125
Db 121 T 121

RESULT 14
US-09-818-875-666/c
: Sequence 666, Application US/09818875
: Publication No. US20030051270A1
: GENERAL INFORMATION:
: APPLICANT: Kmiec, Eric B.
: APPLICANT: Gamper, Howard B.
```

```

: APPLICANT: Rice, Michael C.
: TITLE OF INVENTION: Targeted Chromosomal Genomic Alterations with Modified Single
: TITLE OF INVENTION: Stranded Oligonucleotides
: FILE REFERENCE: Napro-4
: CURRENT APPLICATION NUMBER: US/09/818,875
: CURRENT FILING DATE: 2001-03-27
: PRIOR APPLICATION NUMBER: US 60/192,176
: PRIOR FILING DATE: 2000-03-27
: PRIOR APPLICATION NUMBER: US 60/192,179
: PRIOR FILING DATE: 2000-03-27
: PRIOR APPLICATION NUMBER: US 60/208,538
: PRIOR FILING DATE: 2000-06-01
: PRIOR APPLICATION NUMBER: US 60/244,989
: PRIOR FILING DATE: 2000-10-30
: NUMBER OF SEQ ID NOS: 4385
: SOFTWARE: Friedman macro Napro4
: SEQ ID NO 666
: LENGTH: 121
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-818-875-666

Query Match
Best Local Similarity 100.0%; Score 121; DB 9; Length 121;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 ATTATATCTGCTCTCCGTTGAAGAAGTACAAATGCTATATGCTATGACAGAAATCT 64
Db 121 ATTATATCTGCTCTCCGTTGAAGAAGTACAAATGCTATATGCTATGACAGAAATCT 62

QY 65 TAGAGTGTCCCATCTGCTGAGAGTTGATCAAGAACCTGTCTCCACAAAGTGTGACACA 124
Db 61 TAGAGTGTCCCATCTGCTGAGAGTTGATCAAGAACCTGTCTCCACAAAGTGTGACACA 2

QY 125 T 125
Db 1 T 1

RESULT 15
US-09-818-875-669
: Sequence 669, Application US/09818875
: Publication No. US20030051270A1
: GENERAL INFORMATION:
: APPLICANT: Kmiec, Eric B.
: APPLICANT: Gamper, Howard B.
: APPLICANT: Rice, Michael C.
: TITLE OF INVENTION: Targeted Chromosomal Genomic Alterations with Modified Single
: TITLE OF INVENTION: Stranded Oligonucleotides
: FILE REFERENCE: Napro-4
: CURRENT APPLICATION NUMBER: US/09/818,875
: CURRENT FILING DATE: 2001-03-27
: PRIOR APPLICATION NUMBER: US 60/192,176
: PRIOR FILING DATE: 2000-03-27
: PRIOR APPLICATION NUMBER: US 60/192,179
: PRIOR FILING DATE: 2000-03-27
: PRIOR APPLICATION NUMBER: US 60/208,538
: PRIOR FILING DATE: 2000-06-01
: PRIOR APPLICATION NUMBER: US 60/244,989
: PRIOR FILING DATE: 2000-10-30
: NUMBER OF SEQ ID NOS: 4385
: SOFTWARE: Friedman macro Napro4
: SEQ ID NO 669
: LENGTH: 121
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-818-875-669

Query Match
Best Local Similarity 100.0%; Score 121; DB 9; Length 121;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 AGAAATCTTAGAGTGTCCCATCTGCTGAGAGTTGATCAAGAACCTGTCTCCACAAAGT 115
```



```

Db      1  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          AGAAATCTTAGAGTGTCCCATCTGTGAGTTGATCAAGGAACTGTCTCCAAAGT 60
OY      116 GTGACCATATTTTGCAGATTTGCAACTTCTCAACCGAGAGAAAGGCGCTT 175
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GTGACCATATTTTGCAGATTTGCAACTTCTCAACCGAGAGAAAGGCGCTT 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
OY      176 C 176
          |
Db      121 C 121

```

Search completed: June 27, 2003, 05:25:04  
 Job time : 736.151 secs